DOCUMENT RESUME

ED 201 452 RC 012 680

AUTHOR	Welch, Richard O.
TITLE	Academic Achievement of Mexican American Students: The Edgewood School Plan.
INSTITUTION	D.C. Development Associates, San Antonio, Tex.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Apr 77
GRANT	OEC-0-72-5309
NOTE	83p.; For related documents, see RC 012 677-679.
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Academic Achievement; Affective Behavior; *Bilingual Education; Comparative Analysis; Elementary Secondary Education; *Federal Programs; Improvement Programs; Individualized Instruction; *Mexican American Education; Mexican Americans; *Multicultural Education; Relevance (Education); Sociocultural Patterns; Socioeconomic Influences; Teacher Attitudes
IDENTIFIERS	*Edgewood Independent School District TX; *Experimental Schools Program; Texas (San Antonio)

ABSTRACT

A five-year bilingual, bicultural program, part of the Experimental Schools Program, employed by the Edgewood Independent School District, San Antonio, Texas, was designed to enhance the affective characteristics of the predominantly Mexican American student population, to train teachers to avoid socioeconomic and cultural biases, and to improve the academic achievement of the students. Measurement of the program impact was conducted yearly in language, mathematics, science, social studies, reading, and study skills through a quasi-experimental research design, wherein the treatment schools were matched with comparison schools where students had similar socioeconomic status, affective characteristics, and standardized test achievement scores. Annual comparison scores were collected from both groups over a four-year period, with some significantly higher scores in the treatment groups attributed to particular subjects and grades where treatments were employed. However, the general conclusion was drawn that the experimental program had not substantially improved achievement scores when viewed by school, grade, or subject matter in relation to the Edgewood comparison students that did not participate in the program. (Author/JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED201452

D. C. DEVELOPMENT ASSOCIATES, INC.
MANAGEMENT AND GOVERNMENTAL CONSULTANTS
MILAM BUILDING · SUITE 1400
SAN ANTONIO, TEXAS 78205
512/224 2319

ACADEMIC ACHIEVEMENT OF MEXICAN AMERICAN STUDENTS: THE EDGEWOOD SCHOOL PLAN

by
Dr. Richard O. Welch

Dr. Earl Jones, Director

April 1977

RC 012680

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

HOME OFFICE — 1521 NEW HAMPSHIRE AVENUE, N. W., WASHINGTON, D. C. 20036 • 202/332-5293
DENVER — 180 COOK STREET, SUITE 207, DENVER, COLORADO 80206 • 303/321-2278
HOUSTON — 3220 LOUISIANA, SUITE 203, HOUSTON, TEXAS 77006 • 713/522-2254
LOS ANGELES — 2140 WEST OLYMPIC BLVD., SUITE 534, LOS ANGELES, CALIFORNIA 90006 • 213/487-5740

ACKNOWLEDGMENT

The author appreciates the contributions of the technical reviewers to this report:

Dr. William C. Theimer, Jr., University of the Pacific, Stockton, California

Dr. Ernest Bernal, University of Texas at San Antonio

Dr. Milton Goldsamt, Development Associates.

Through the years of data collection and analyses, the guidance of Dr. Robert A. Cervantes, then director of the San Antonio office of Development Associates and now with the Division of Research and Evaluation at the California State Department of Education, and of Dr. Earl Jones, Development Associates, contributed significantly to this report.

Ms. Jayne Wieding typed the preliminary and final copies of this document.

* * * * * * *

The collection of data presented and preparation of the content reported herein were supported by grant OEC-0-72-5309 from the National Institute of Education, Experimental Schools Program, of the Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position nor policy of the National Institute of Education and no official endorsement should be inferred.

TABLE OF CONTENTS

Section	Page
ACKNOWLEDGMENT	i
LIST OF TABLES	iii
ABSTRACT	iv
INTRODUCTION	1
The School Treatments	2
The Achievement Literature	4
RESEARCH DESIGN	8
Research Questions	9
Instrumentation	10
Sampling Plan	11
Data Analyses	13
RESEARCH FINDINGS	16
First Year Findings	16
Second Year Findings	18
Third Year Findings	22
Fourth Year Findings	33
SUMMARY AND DISCUSSION	36
First Project Year	36
Second Project Year	36
Third Project Year	37
Fourth Project Year	38
Recapitulation	39
BIBLIOGRAPHY	42
APPENDICES A & B: SUMMARY TABLES OF SIGNIFICANT RESULTS: 1972-1973, 1973-1974	46
APPENDIX C: SUMMARY TABLES OF SIGNIFICANT RESULTS: 1974-1975	49
APPENDIX D: SUMMARY TABLES OF SIGNIFICANT RESULTS: 1975-1976	63

LIST OF TABLES

Table		Page
1:	Number of ESP and Comparison Students Tested with the TOBE and CTBS/Q: 1972-1973 through 1975-1976	12
2:	Mean Achievement Raw Scores: 1972-1973	17
3:	Percentage of Students Scoring Below National Norms: TOBE and CTBS/Q Subtests, 1972-1973	18
4:	TOBE/CTBS/Q Raw Score Means by Grade and School: 1973-1974	19
5:	Mean Raw Scores of Elementary and Secondary Students Analyzed from the First to the Second Project Year	22
6:	Summary and Direction of Findings by Group and Sex: Grades K-6, 1974-1975, 1975-1976	24
7:	Summary and Direction of Findings by Group and Sex: Grades 7-12, 1974-1975, 1975-1976	25
8:	Absolute Differences between Statistically Significant Treatment and Comparison Raw Score Means, CTBS/Q: 1974-1975	28
9:	Eta Squared Coefficients on CTBS/Q Total Battery, Grades 8, 10, 11, and 12: 1974-1975	31
10:	Summary of Significant Differences between ESP and Comparison Mean Scores by Quartile, TOBE-CTBS/Q: 1974-1975	33
11:	Percentage of Students Scoring below National Norms: TOBE and CTBS/Q Subtests, 1975-1976	35

LIST OF FIGURES

Figure		
1:	Kindergarten TOBE Means by Group and Sex: 1974-1975	26
2:	First Grade TOBE Means by Group and Sex: 1974-1975	27
3:	Mean Raw Scores on CTBS/Q Total Battery by Group and Sex, Grades 8, 10, 11, and 12: 1974-1975	30

ABSTRACT

Academic achievement was the ultimate goal of the Edgewood School Plan, a National Institute of Education experimental schools program in the Edgewood Independent School District in San Antonio, Texas. Locally developed and administered intervention programs were implemented for the largely Mexican American student body in four elementary, one middle, and one secondary schools designated to receive the experimental treatment. These schools were chosen from the twenty-six in the district on the basis of constituency, a mini-system of feeder schools K-12. Two elementary, one middle, and one secondary schools were selected as comparison schools on the basis of similar ethnic group composition, socioeconomic level, achievement, and affective characteristics as the treatment schools. Five year federal funding was provided for the project and Development Associates contracted to do the external evaluation of the first four of the five years.

Each year the Tests of Basic Experiences were used as measures of achievement in grades K and 1; the appropriate levels of the Comprehensive Tests of Basic Skills, form Q, were administered to grades 2-12. In 1971-1972, the baseline year, the experimental and comparison groups scored alike except on one subtest. No important differences existed between these groups.

In the second year, the treatment kindergarten and first grade children scored significantly higher on all the subtests than did those in the comparison schools. Fifth graders in the experimental schools scored higher in language achievement than did comparison children. The treatment secondary students scored significantly higher on all subtests and the total battery than those in the comparison high school. A large increase in mathematics accounted for most of the total battery score difference. The gains in mathematics from the first to the second year were greater for experimental than comparison. During this year treatment programs had been in effect in one of the experimental schools' K and first grades, in the language instruction in their fifth grades, in the secondary mathematics and in most of the other ninth grade classes. The differences between experimental and comparison were not uniform across grades or schools except as specified, however.

In the third year, the experimental kindergarten and first grades continued to be significantly higher in achievement than comparison. For grades 2 through 7, however, the only significant differences were in favor of the comparison students. Experimental grades 8 through 12 continued to exhibit higher scores in some subtests than comparison. The research design examined the role of sex factors. Several of the significant differences were accompanied by significantly higher scores for females, especially those in the experimental program. An examination of the treatment curricula gave no indications that the thrust of the experimental programs favored one sex over another.

The fourth year achievement data showed the experimental kindergarten and first grade students continuing higher in mathematics than comparison students but not in language. Treatment second graders were higher on all subtests and the total battery. There were no differences at third grade between the two groups. There were several statistically different scores at fourth through sixth grades, all in favor of the comparison students. Seventh grade data showed no differences by group. Isolated differences in favor of comparison occurred at the eighth and ninth grades. Scattered differences with experimental higher were found at the other grades. Although fewer than for the previous year, a number of sex differences, all with females higher, remained during the fourth year, 1975-1976. In many cases throughout the K-12 span, the treatment students scored lower during the fourth year than the corresponding grades had scored the year before. The several elementary schools still showed significantly different means, with one comparison school highest and some treatment schools decidedly lower than this school. An examination of the programs in these schools did not provide a consistent explanation for the differences.

Several of the data patterns evident in previous years were interrupted in 1975-1976. Some of the statistical differences occurred, or were not repeated, because the experimental students' scores were lower than the previous year, rather than that the comparison scores had changed. Relatively massive individualized instruction treatments were infused into most classes in the experimental schools in the fourth year. The comparison high school also introduced substantial individualized curriculum packets into its program. Teachers and students may have experienced adjustment difficulties with these efforts that detracted, at least temporarily, from the general progress.

Examination of the cross sectional achievement data across the four years provided important general information:

1. Neither the treatment nor the comparison group changed noticeably over the four years from the pattern of becoming increasingly further behind the national norm as the students progressed through grades K-12.
2. Performance in the elementary schools was not uniform; neither the experimental nor the regular instructional program produced consistent results across schools.
3. The achievement means of the treatment and comparison groups did not remain stable across the four years; the comparison group showed some losses the third year, experimental the fourth.
4. Performance was not uniform between girls and boys; the experimental females were frequently and often substantially higher than experimental males and either sex in the comparison schools.

5. No convincingly consistent pattern of achievement differences was encountered when the scores were analyzed by performance quartiles or when compared to national norms.

When these generally inconsistent patterns were taken into account, the conclusion must be that the experimental program was unable to produce general academic achievement increases with its treatments.

In brief, the Edgewood School Plan, during its first four years of funding, introduced a variety of treatments over different grades and for varying lengths of time. Some consistency of treatment was observed in the general approach and in the staff training supplied to teachers. When viewed annually, some significantly higher experimental students' scores appeared to be associated with treatments in the particular subjects and grades where differences between experimental and comparison groups were noted. The experimental program, however, tended to change its treatments often or at least to modify them considerably; persistence of treatment effect was not encountered. In the aggregate, the considerable inconsistencies led to the general conclusion that the experimental program had not substantially improved the achievement scores of its pupils when viewed by school, grade, or subject matter in relation to the Edgewood comparison students that did not participate in the program.

INTRODUCTION

Differential school achievement among ethnic groups is a cause for serious concern because school performance has become the chief avenue for upward social mobility in American society. When, with few individual exceptions, an entire group of people distinguished by common socioeconomic or ethnic traits consistently achieves below other groups characterized differently on these traits, it means that the group's lifetime opportunities are being systematically depressed. That such a situation prevails for the Mexican American has been extensively documented. (9, 29, 34) Locating the cause of this underachievement is more difficult.

Early studies of this phenomenon elsewhere delineated and correlated characteristics of the Mexican American culture which were thought to have adverse effects on achievement. Many recent studies reported in the literature have rejected such approaches, however, and have ascribed the failure to the educational institution's general inability to adapt programs to the distinctive needs and characteristics of Mexican Americans or to inadequate educational programs and facilities resulting from the Mexican Americans' marginal socioeconomic position. (1, 29)

The present research measured the impact of a large scale bilingual-bicultural program, following the trend of these later studies and designed on the premise that appropriate changes in the educational institution would facilitate increased, competitive achievement among the Mexican American students participating in it. The specific program investigated was the Edgewood School Plan (ESP) of the Edgewood Independent School District (EISD) in San Antonio, Texas, an educational change program with five year longitudinal federal funding by the National Institute of Education (NIE). This plan proposed to alter various educational processes to create "compatibility" between the educational institution and Mexican American students' characteristics and needs. In particular, ESP proposed to develop an educational program which would promote the psychos ial growth of Mexican American students, including the improvement of their self concept, motivation, and school attitudes, and thereby increase their school achievement. (14)

The external evaluation of the project was contracted by the National Institute of Education with a private company, Development Associates, which set up a San Antonio office to collect and analyze the data for the summative study. Utilizing the goals and objectives stated in the ESP Final Plan (14), and modifying the instrumentation to adapt to programmatic developments, the researchers monitored the processes of project planning, management, and implementation.

They also measured the affective variables of self concept, motivation, and attitudes, and the academic achievement of the pupils in the project's target grades. The results were analyzed for each year and then over the four year period.

The present report is one of a series produced by Development Associates in the full documentation of the Edgewood School Plan. The Second Interim Report (7) described the first two years' activities and presented preliminary analyses of the affective and cognitive data. During the next two years, affect was also treated in specialized reports on self concept, motivation, attitudes, values, and locus of control (5, 7, 8, 26, 32, 35), and was also combined with achievement to study their interrelationships in a four year investigation (6). The present report concerns itself only with achievement and examines the data grade by grade, on a year to year basis, to determine short term (each year) differentials by group (treatment, comparison), subjects tested (reading, language, arithmetic, and study skills), and sex. Special analyses by level of performance also are included.

The School Treatments

This research involved both a treatment (ESP) and a comparison group. The treatment schools included four elementary, one middle, and one high school. The comparison group was comprised of two elementary, one middle, and one high school, selected from among the twenty-six in the district that were most like the treatment schools.

The treatment project was called the Edgewood School Plan and primarily stressed two student improvement areas: affective and cognitive. Both of these were addressed in the Final Plan (14) and in several subsequent ESP documents, especially in the Philosophy of Communication Skills (16) and Personalizing Instruction Models and Program Emphasis (15). Essentially, the affective programs were directed toward improving teacher pupil relationships through teacher workshops and conferences, and counseling. Some efforts toward increasing the relevancy of curricular materials through the inclusion of Mexican American and local Texas topics were evidenced, particularly during the first year. The use of Spanish in some classroom management and teaching was included as a means of affective enhancement.

In the cognitive domain, most of the early elementary curricular materials produced by ESP were supplementary, bilingual, and bicultural. A special bilingual integrated subject matter approach program was developed for kindergarten in one school and was fully implemented in that school only during the second year. It was partially implemented in the third year. Diagnostic reading and individualized instruction in mathematics were utilized in some of the ESP fifth grade classrooms the second year.

Some training for implementation was provided. During the first three years, these training efforts were mostly directed toward grades K-2. Several commercial materials were made available for bilingual education. The implementation of these appeared to be at the school's and/or teacher's option. In the fourth year, an ESP spiral mathematics program, an individualized, contract approach language arts project, and some social studies instruction were provided for K-6.

During the first project year some teacher-staff produced and commerical materials were utilized at the middle school level with some team teaching. During the second and third years, the materials and the team teaching procedures were revised and included mathematics, science, humanities, social studies, and language arts. These were further refined for the fourth year and the mathematics program, previously arranged in LAP* sequence, was changed to a spiral arrangement.

The beginning secondary materials were both teacher-staff produced and commerical, and were taught to limited numbers of students. The second year, the LAP sequence mathematics program was used for most middle school students and all students in the secondary classes; this, too, was revised to the spiral arrangement for the fourth year. Social studies, language arts, and some other subjects were taught to a portion of ninth grade, using teacher-staff produced materials after the first year. These underwent revision through the project's life and were subsequently expanded to tenth grade.

The comparison elementary schools used a general bilingual education approach for the Spanish dominant and bilingual English-Spanish speaking children in grades K-2. Most of the other elementary and middle school instruction was implemented through traditional textbooks and was taught in English. The comparison high school was generally traditional in its materials and methods during the first three years of this study. In the fourth year, however, a specially funded program facilitated the preparation of teaching/learning units in English, social studies, and science and these were offered through individualized instruction. Commercial materials in mathematics and reading were purchased and taught in a self paced, individualized program. About 75% of the students in the comparison high school was included in this project. (24)

*LAP = learning activity package. Objectives are arranged in sequence from entry to terminal objectives for each unit or segment of learning.

In brief, there were diverse sets of specific and general affective, cognitive, and behavioral program treatments directed at students in the elementary, middle, and high school levels, as well as a series of treatments directed at improving teacher competencies. Detailed specification of these treatments and other related and community work conducted by ESP during its first two years of operation was included in the Second Interim Report (7). Further studies of these areas and of those treatments initiated during the third and fourth years of the program are currently in preparation. (22)

The Achievement Literature

The ESP goals arose from research studies which had suggested that educational institutions had not adapted themselves to Mexican Americans' needs and characteristics, or had inadequate features. Thus, adverse effects on achievement had resulted. Much of the research literature on the educational level of Mexican American students had consistently reported academic achievement below the levels of national norms for other White and some minority groups. The causes cited for the low achievement level generally have been related to socioeconomic status, and pedagogical and institutional practices.

The first major empirical account of the educational status of minority groups was that by Coleman et al (9), Equality of Educational Opportunity, published in 1966. Although this study, commissioned as part of the 1964 Civil Rights Act, was concerned primarily with ascertaining the extent of educational segregation and levels of ecuational opportunity of Blacks, data were compiled on other minority groups, including Mexican Americans. The study concluded that the educational achievement of Mexican Americans was below that of other White and Black students but above that of Puerto Ricans and American Indians. The most pronounced deficiencies were in mathematics, reading comprehension, and verbal ability. By the twelth grade, Mexican Americans as a group scored 4.1 years below the national norm in math achievement, 3.5 years in verbal ability, and 3.3 years in reading.

The most extensive study on Mexican Americans was that conducted by Grebler, Moore & Guzman (20), who studied economic, social, political, and educational factors. The authors noted that the median educational level in the mid 1960s ranged from a low of 4.8 in Texas to a high of 8.6 in California, as compared to 11.5 and 12.2 for Anglos. Reading and verbal skills were the areas in which the most serious deficiencies appeared. They also noted that cultural diversity, language, educational exposure, and socioeconomic factors among Mexican Americans strongly affected the latter's achievement.

The most recent reports documenting the educational status of Mexican Americans were contained in a series produced by the

United States Commission on Civil Rights. Although they suffered from some methodological weaknesses, the major contribution of this series was a greater understanding of some of the causes of low educational attainment. The commission's report, The Unfinished Education (36), concluded that without exception, Mexican American students achieved at a lower rate than other White students, school holding power was lower, reading achievement was poorer, grade repetition more frequent, overageness more prevalent, and participation in extra curricular activities less than for their Anglo counterparts. The report estimated that out of every 100 Mexican American pupils who entered first grade, only about 60 graduated from high school. It was also estimated that 40% of the Mexican American students in the Southwest, between 8th and 12th grades, was reading two or more years below grade level. Another report, The Excluded Student (37), concluded that the low educational achievement was in part attributable to the tendency of schools to ignore the language, culture, and heritage of students; the over representation of Mexican Americans in low ability groups and educable mentally retarded classes; and the under representation of Mexican American school staff. A later report, Mexican American Education in Texas: A Function of Wealth (38), examined the manner in which the Texas school finance system discriminated against schools with high proportions of low income groups, including Mexican Americans.

In an investigation of the trend of educational status of Mexican Americans in Texas between 1950 and 1960, Lamana & Samora (29) concluded that the median level of education had increased from 3.6 to 4.8 years but that the median was still 1.2 years below that of the total population; the educational status of Mexican Americans in Texas was the lowest in the Southwest; that "better" educated Mexican Americans were becoming geographically segregated from the less educated; and that the educational status of Mexican Americans relative to other groups in Texas was unchanged from 1950 to 1960.

Development Associates had analyzed other data on the same schools included in the present study. The findings for the first two years of the Edgewood study were reviewed in two interim reports (7, 13) and a four year study of achievement was included in a later publication (6). The analyses included in the three reports showed little change in the standardized achievement scores of the students in these schools. One general trend, however, was for girls to score higher than boys. This was consistent with the findings of some other research on Mexican Americans (5, 12, 17, 18).

Other studies carried out by Development Associates in 1975 and 1976 indicated that, for the most part, the Mexican American students tested in two school districts in South Texas scored below the national norms. A number of students in one district were tested in grades 1 through 3 during 1975, and grades 1 through 12 in a pre - post evaluation design during

1976. The 1976 post test revealed that the first graders scored .3 of a year below norm and the eighth graders scored 1.1 years below the norm on the total battery of achievement tests. (27)

In an evaluation of one school district's ESAA projects during the 1975-1976 school year, Development Associates found that although the students' mean gains were more than one year, their achievement scores at the end of the year were still below norm (25). The scores of fifth, sixth, and seventh graders in mathematics fell 1.1, 1.6, and 1.8 years below norm, respectively. Reading scores for the sixth and seventh graders were even lower and gains were less.

A major exception to Mexican American students scoring below national norms in South Texas was that found in another small, predominantly Mexican American district. (23) The results from grades 1, 2, and 3 showed that, in general, achievement was at the norm levels or above for the district students. The remaining grades, 4-12, scored very near the norm. Some of the success of these results was attributed to the effects of an existing tradition of students going to college, graduating, and making significant contributions to the professional world. And to continue this tradition, the project staff was not satisfied with "average" achievement in the primary grades - it worked to assure mastery of the basics to enable every student to profit from the total academic program. Revised curricula, staff development, and improved methods were believed to have accomplished the norm achievement over a five year period.

In brief, there is considerable evidence documenting low educational achievement of many Mexican American students. The major reasons cited in the research literature were the effects of poverty, discriminatory educational practices, failure to positively recognize the language, culture, and heritage of students, and under representation of Mexican American faculty. The accumulated effect of these school practices was that Mexican American students were provided an inferior education. Districts that have overcome much of the problem have cited increased attention to the basic skills as the enabling factor.

Much of the contemporary literature is particularly critical of the socio-cultural assumptions often cited as the cause of low achievement and overt discriminatory practices. In their investigation of socio-cultural determinants of achievement among 263 Mexican American children, Anderson & Williams (1) found that these children desired good grades and received as much encouragement at home as Anglo children. The authors further concluded that the consistently low achievement of Mexican Americans was due to inadequate educational programs and not to low motivation. Carter (4) also cited the lack of adequate provisions in school programs to meet the needs of

Mexican American students as accounting for low achievement. He further concluded that the stereotypic notion of "cultural deprivation" was primarily a rationalization to excuse the weak development of positive educational programs. In a discussion of the educational problems facing Mexican Americans, Ortego (34) also charged that inadequate school programs and continued use of socially and educationally discriminatory practices were at fault for low educational achievement.

With regard to the effects of particular schools, Mayeske (31) stated that about 20% of the difference in educational achievement among Mexican Americans was attributable to the kinds of schools they attended. Although Grebler (19) found that there was evidence of improvement in educational achievement of both groups between 1950 and 1960, the gap between Mexican American and other White children remained substantial. Grebler attributed the educational lag to the late entry of Mexican American children into formal education, differential proportions of school enrollment between Mexican Americans and other Whites, and the low percentage of Mexican Americans graduating from high school.

More recent research studies by Lorea & Kohut (30), Mycue (33), Del Buono (12), and Broussard (2), all of which controlled for linguistic, socioeconomic, and cultural variations, concluded that Mexican Americans were not atypical of other minority students. Darcy (11) noted, based on the review of a decade of research, that bilingual children were not inferior to monolinguals when non-verbal tests were used and socioeconomic status was controlled. Similarly, Gezi (18), on the basis of his review of research on bilingual-bicultural education, found that bilinguals had IQs as high as those of monolinguals.

In general, research on achievement focusing on Mexican American students has documented below grade performance and, in many other cases, an inferior quality of education. The general pattern is a widening gap in achievement scores, when compared to other groups, as these students progress through the public schools. Although various causes were cited for low achievement performance, most investigators agreed that failure to address one or several of the following factors contributed to it: social, pedagogical, and linguistic needs of students, or failure by educational institutions to positively take students' cultural attributes into account in the instructional process. Some studies which investigated Spanish language instructional methodologies and controlled for socioeconomic status suggested that Mexican Americans often attain as high achievement scores as do other White students.

In addition to the causes cited for the low achievement of Mexican Americans, it is probable that the national trend in achievement scores has also affected these students. Since the mid 1960s, test score means have dropped. In a recent

article, Harnischfeger & Wiley documented this decline as follows (21, p. 10):

> Beyond doubt, and beyond differences among assessment instruments, alterations in tests, or in pupil composition, achievement test scores have been declining for about a decade in all grades from five upwards. Score declines are more pronounced in higher grades and in recent years. They are more severe for tests probing verbal than mathematics achievement. These are the facts and they describe a national phenomenon.

While the article did not include statistical analyses, the authors had studied a large amount of data from a number of standardized tests and for students from most geographical areas. Trends in data were followed systematically from 1972 onward.

RESEARCH DESIGN

The Edgewood Final Plan was written to enhance the affective characteristics of the mostly Mexican American students, to train teachers to avoid socioeconomic and cultural biases, and ultimately, to improve the academic achievement of these students.

The measurement of student academic achievement to assess the treatment impact of the Edgewood School Plan focused on an examination of student performance in language, mathematics, science, and social studies at grades K and 1, and in reading, language, arithmetic, and study skills at grades 2 through 12 over the four year period of the program. These variables were examined through a quasi-experimental research design, that is, students within intact grades were tested (3). The ESP (treatment) schools were matched with others (comparison) from the same district that had approximately the same composition of ethnic groups, and where the students were about alike in socioeconomic status, affective characteristics, and standardized test achievement scores.

The students in each group were tested with an achievement battery and the mean scores at all applicable grades were analyzed for differences each year. The program planners expected that the mean scores of the ESP students would be significantly higher than the mean scores of the comparison students in the grades where ESP treatments were carried out.

This approach allowed for continuity in the reporting of student achievement data in both the treatment and comparison schools. Each year the ESP project staff designated certain grades for treatment. These treatments provided the framework

for data interpretation. The mean scores of the treatment and comparison students were also analyzed with respect to national norms on the appropriate tests to assess relative shifts in either group over the four years of the program.

Research Questions

One major research question guided the analysis of the achievement data collected throughout the project period:

1. Did the ESP students obtain significantly different achievement scores than the comparison students?

A corollary question was:

2. Were there significant sex differences in achievement in the treatment or comparison groups?

To help in clarifying results based on the major research questions, supplementary analyses were carried out dealing with the following questions:

3. What were the year to year gains in the achievement scores from the first to the second year?
4. Did the program benefit students of higher or lower educational performance levels during the third year?
5. Did the treatment and/or comparison students score closer to the national norms after four years of project treatments?

Data were analyzed each year at all grades for which achievement scores were collected from both ESP and comparison students. As shown in Table 1, these grades included K, 1, 5, 8, and 10 during the first year; K through 6, 9, and 11 during the second year; and K through 12 during the last two years.

Development Associates, to determine comparability between the treatment and comparison groups the first year (1972-1973), used analysis of variance techniques to determine if significant differences existed between the mean scores of the achievement subtests and total battery for the students in each group. Virtually no differences were found for that year.

A further aspect of the first research question involved examining the achievement scores across the six ESP and comparison elementary schools. If scores of the ESP students were found to be significantly different than those of the comparison students at one of the elementary grades during the second, third, or fourth year, it was necessary to determine if the difference was due to students in particular ESP schools. Since the treatments carried out in the four ESP elementary schools were not always identical, even at the same grade, the possibility of differential results existed.

The second research question pertained to an investigation of the achievement scores by sex. Although the project did not intentionally focus the treatments on boys or girls separately, it was possible that one sex or the other would be more receptive to the methods and curricular content.

Although data on student sex were not collected during the first two years for reasons beyond the control of the researchers, such data were available for the last two years and analyses were performed at each grade for these years. As part of the investigation of the achievement scores, Development Associates also examined the interaction effects between group (treatment, comparison) and sex.

Instrumentation

The Tests of Basic Experiences (TOBE) and Comprehensive Tests of Basic Skills, Form Q (CTBS/Q), were administered each spring from 1973 to 1976.* The TOBE was administered to kindergarten and first grade students, and the CTBS/Q to second through twelfth grade students.

The TOBE consisted of two forms, with level K used for kindergarten, and L for the first grade. The TOBE tests the areas of mathematics, language, science, and social studies.

The CTBS/Q consists of test levels specific to certain grades: level 1 for grades 2, 3, and 4; level 2 for 5 and 6; level 3 for 7 and 8; and level 4 for grades 9 through 12. The CTBS/Q tests the areas of reading (vocabulary, comprehension), language (mechanics, expression, spelling), arithmetic (computation, conceptualization, application), and study skills (reference skills, use of graphic materials).

Both the TOBE and the CTBS/Q have established adequate reliability and validity. In terms of the Kuder-Richardson formula for measuring internal consistency (KR_{20}), the reliability of the TOBE ranged between .73 and .84 for the kindergarten and first grades. For the CTBS/Q, the KR_{20} coefficients ranged between .75 and .96 for across grade levels (2 through 12) with most of the coefficients being over .80. Several studies have been carried out to establish the content validity and criterion related validity of the TOBE and CTBS/Q (10). A detailed discussion of the reliability, validity, norms, and other

*The ESP staff also tested some grades in fall 1975. Development Associates did not use those scores in any of their analyses.

characteristics of the TOBE and CTBS/Q were presented in the First Interim Report (13).

The investigators cooperated with both the EISD Office of Pupil Assessment and the ESP internal evaluation unit in the administration of the achievement batteries each spring. In-service training was provided to the teachers for administering the tests in the respective classrooms. The tests were given to the treatment and comparison students during the same month, and testing was carried out within approximately the same two week period, thus removing possible time influences from the obtained results.

For the most part, the TOBE and CTBS/Q tests were scored by a local computer facility. The Level 1 booklets of the CTBS/Q covering grades 2, 3, and 4, were scored by the test publisher in California during the first two project years since these forms required special scanning equipment. These were scanned at a computer facility in San Antonio thereafter. The printouts provided raw scores for each student on all the subtests and either stanines (TOBE) or grade equivalencies (CTBS/Q) in terms of national norms. Also the mean raw scores and mean grade equivalencies were calculated for each elementary school classroom and each middle and high school grade tested.

Sampling Plan

Table 1 depicts the number of students at each grade in the ESP and comparison schools who were tested with the TOBE and CTBS/Q series throughout the four years of the project. The blanks for some grades indicate where students were not tested as a result of some modifications in the sampling plan. The total number tested ranged from 4,660 in the first year, or about 50% of the students enrolled in the ESP and comparison schools at the time of testing, to 7,526 in the third year, or about 85% of the students enrolled in the schools at that time. Absentees accounted for all those not tested in 1974-1975 and 1975-1976. The absences were greater at the middle and secondary levels, approaching 20%, and lower, generally about 10%, at the elementary level. No important difference in the proportion absent during the testing period occurred between the treatment and comparison schools.

Because several groups were cooperating in the collection of achievement data, the sampling approach was modified from year to year over the duration of the project to accommodate the needs of each group. For example, although ESP had the primary responsibility for the achievement testing in all four years, the project staff did not have an internal evaluation unit during the first year and Development Associates carried out the testing under the auspices of NIE. At that time, it was decided to test all students in intact classes from grades 1 through 11 in the ESP schools, to randomly select one intact

kindergarten class from each of the ESP schools together with four from the comparison schools, and to randomly select four intact first, fifth, eighth, and tenth grade classes from the comparison schools for testing. The twelfth grade students in both the ESP and comparison schools were not tested since most of these students had already completed their class work for the year and were preparing to graduate. The other grades not tested in the comparison schools that year had been tested the previous month by the district with another achievement instrument. It was believed that "practice effects" might appear in results if these students were included.

Table 1: Number of ESP and Comparison Students Tested with the TOBE and CTBS/Q: 1972-1973 through 1975-1976

Grade	1972-1973		1973-1974		1974-1975		1975-1976	
	ESP	Comp	ESP	Comp	ESP	Comp	ESP	Comp
K	124	124	*300	157	*275	172	*326	146
1	400	123	*355	201	*359	191	*332	201
2	426	-	353	185	*365	182	*350	173
3	440	-	380	195	378	179	*324	174
4	413	-	385	219	392	185	*353	168
5	308	128	433	221	419	200	*375	187
6	406	-	352	179	369	206	*396	177
7	354	-	*367	-	*268	339	*285	347
8	316	105	*340	-	*315	297	*299	300
9	345	-	401	488	*461	420	*491	400
10	304	94	-	-	288	360	*402	386
11	250	-	217	334	181	260	216	266
12	-	-	-	-	197	268	177	185
Total	4086	574	3883	2179	4267	3259	4326	3110
Total ESP and Comp	4660		6062		7526		7436	

*Indicates ESP designated treatment grade.
NOTE: Comp=comparison; ESP=treatment.

During the second year, all students in intact classes were tested except in grades 7 and 8 in the comparison schools. The comparison seventh and eighth graders were dropped from the testing program for reasons beyond the control of the researchers. Grades 10 and 12 were tested that year by the district with a different achievement instrument and were not retested with the TOBE-CTBS/Q series in order to avoid practice effects.

All students in intact classes were tested in all grades during the last two years of the program. In the last year, however, the kindergartners and first graders in the two comparison schools were tested only with the mathematics and language

subtests of the TOBE to conform with other testing activities carried out in the district.

Although this sampling plan was adequate for the cross sectional analyses of this study, the problems of tracking all students across the four years of the project proved difficult. Student attrition due to dropouts, transfers, and the district feeder system modifications made it impossible to follow all of the students in an analysis of the longitudinal samples. That portion of the Development Associates evaluation design was reported separately (39) and is not included in the present report. In another study, analyses of cross sectional, longitudinal, and repeated measures were performed on the students tracked across the four years (6). Those students were divided into three cohort groups and several affective as well as cognitive variables were examined.

Data Analyses

The sampling plan enabled several analyses to be carried out each year. In the first program year the ESP project staff spent most of its time planning and organizing, and initiated treatments only in late spring, thus minimally affecting achievement. Consequently, Development Associates collected the first year data as baseline information which was employed as a "before" project measure. As treatments were initiated in the following years, these data were then used as a comparison basis for assessing the student progress in achievement.

The first year data were analyzed for all grades for which achievement scores existed for both the treatment and comparison students for grades K, 1, 5, 8, and 10. Two analyses were carried out:

1. to determine if one group were significantly different than the other before the project treatments began as a basis for explaining subsequent results, and
2. a supplementary analysis to determine how each group scored vis-a-vis the national norms.

For the first analysis, a sample of 60 students (30 from each group) was randomly selected from the grades tested and one way analysis of variance tests were performed to examine between group differences. The alpha level for determining statistical significance was set at .05 for all tests carried out that year and in subsequent years. For the supplementary analysis, all student scores from the grades tested were examined and compared to the national norms.

For the second year, data were collected for grades 1 through 6, 9, and 11. For the first research question the classroom

means for each school at each elementary grade were examined in a series of between school, one way analysis of variance tests. When significant differences were found, post hoc multiple comparisons were carried out on all pairs of school means using the Scheffe procedure to determine where the differences existed among the ESP and comparison schools. At the secondary level, the data for the treatment and comparison ninth and eleventh graders were analyzed in a 2 by 2 factorial design. If significant differences were found by group (treatment vs. comparison), then post hoc multiple comparisons again were performed to determine where the differences existed.

In a supplementary study during the second year, an analysis of covariance technique was used to determine if treatment and comparison students differed, relative to their initial levels of achievement, during the first year. (In general, analysis of covariance has been considered by researchers to be superior to computing raw difference scores as a means of determining achievement gains.) To maximize the use of existing data, the scores of two groups of treatment and comparison students were analyzed for the two years - one group at the elementary level, the other consisting of a group of high school students. At the elementary level, the second year achievement scores of all treatment group sixth graders were compared with the scores of the same students who took the battery as fifth graders. These changes were then tested against the changes in the second year achievement scores of the comparison group sixth graders who were administered the battery as fifth graders. A similar within groups analysis of covariance was performed using the treatment and comparison group tenth and eleventh graders and evaluating the trend in data from the first to the second year.

The collection and analysis of achievement data in the third year were more comprehensive than for the previous years. As shown in Table 1, all students in all grades were tested in both the treatment and comparison schools. Furthermore, both the research questions dealing with the effects of the treatments by group and sex were examined using a 2 by 2 factorial design at each grade. If significant differences were found by group at the elementary level, the data were reanalyzed by school, with post hoc multiple comparisons computed for all pairs of means to determine where the differences existed among the six treatment and comparison schools.

Since the scores of all students tested were investigated that year, two additional procedures were employed to make certain that the data used in the ANOVA analyses were representative of and consistent with proper statistical design. It was felt that using these procedures would strengthen the statistical precision of the analyses. The first procedure involved eliminating the students who:

1. did not complete an entire subtest on either the TOBE or CTBS/Q, and
2. could not be identified by sex.

Most of the schools administered the tests over a two day period, and if some students were absent during one of these days, they did not complete all of the subtests. Consequently, those students were eliminated from the analyses. The students who could not be identified by sex were eliminated since this variable was necessary to the factorial design.

The second procedure employed the third year was to equalize the cell size between the two groups of treatment and comparison students at each grade. This was particularly important at the elementary level where the students tested in the four treatment schools outnumbered those tested in the two comparison schools. At the junior and high school levels, the case was sometimes reversed with the comparison students tested outnumbering the treatment students. Furthermore, at both the elementary and secondary level, the number of students tested in one group often was not proportional to those tested in the other group. A random sample of students was selected from the larger group tested to match the number of students tested in the smaller group at each grade.

The scores of all students not included in each grade for the 2 by 2 factorial analysis for any reason were analyzed separately. The means of these scores were computed and compared to the means of the group from which they were eliminated. In all cases, these comparisons revealed that neither the mean nor the variance of the scores of the students not selected (from the process of equalizing the N sizes) differed significantly from the mean or variance of the scores of those selected.

In a supplementary analysis carried out during the third year, the question of whether the program benefitted students of higher or lower educational performance was explored. Performance was defined in terms of the distribution of scores by quartile so that those students scoring in the top two quarters (lowest scores) were grouped as the "low" performers while those scoring in the bottom two quarters (highest scores) were used as the "high" performing group.* As with the question related to sex, the project staff did not focus the treatments on either the students of higher or lower performance separately, but one group or the other might have been more receptive to the methods and curricular content.

*The "high" and "low" groups did not divide the number of students exactly at 50% since whole number scores precluded that division. The variation was small, however, between the cell sizes for any analysis.

Grades K, 1, 8, and 10 were selected for this analysis since the students in these grades received most of the ESP emphasis during this year. Also, using these grades yielded a sample of all three levels of schooling: elementary, middle, and secondary. At each grade, t tests were performed on the mean scores of all TOBE or CTBS/Q subtests at each quartile between the treatment and comparison students. The pattern of results was then analyzed by grade to determine if there were differences.

The analysis of the data during the fourth year consisted of a duplication of the analyses performed in the third year to determine if significant differences existed between the treatment and comparison groups, controlling for sex (research questions one and two), together with a followup to the analysis of the first year to determine how the two groups scored in relation to the national norms after the four years of program exposure. As in the third year, a 2 by 2 factorial design was used to analyze the scores by group and sex at each grade. Also, the same two procedures were employed whereby students were eliminated who either did not complete a subtest on the achievement battery or could not be identified by sex. The number of the treatment and comparison students was equalized at each grade.

RESEARCH FINDINGS

Each year's findings from the data analyses are presented separately. As pointed out in the previous section, the first research question was investigated utilizing the data from all four years of the project. The second research question was analyzed using the third and fourth year data. The findings from the supplementary analyses are discussed under the year to which they apply.

First Year Findings

In regard to the first research question, the treatment students did not obtain significantly different achievement scores from the comparison students during the first year of the program. Thus it was felt that the sets of students could be compared for subsequent years. Table 2 shows the means and cell sizes for the student scores analyzed in both groups on all the achievement subtests for that year.

The results of the between group, one way ANOVA comparisons performed on these means were presented in the Second Interim Report (7). These demonstrated that only one significant difference existed between the treatment and comparison means and that difference occurred on the arithmetic subtest at the

eighth grade. In that case, the mean of the comparison students (51.57) was significantly higher than the mean of the treatment students (42.83). A summary of the ANOVA results obtained by using this subtest is shown in Appendix A1.

Table 2: Mean Achievement Raw Scores: 1972-1973

		Treatment		Comparison		Total	
Grade/TOBE	Subtest	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
K:	Language	20.13	30	22.00	30	21.07	60
	Mathematics	19.03	30	19.60	30	19.32	60
	Science	19.07	30	19.57	30	19.32	60
	Soc. Studies	19.13	30	18.77	30	18.95	60
Grade/TOBE							
1:	Language	17.23	30	19.73	30	18.48	60
	Mathematics	17.20	30	18.80	30	18.00	60
	Science	14.50	30	16.87	30	15.68	60
	Soc. Studies	16.47	30	18.77	30	17.62	60
Grade/CTBS/Q Test							
5:	Reading	35.80	30	34.63	30	35.22	60
	Language	39.90	30	36.50	30	38.20	60
	Arithmetic	52.87	30	48.63	30	50.75	60
	Study Skills	20.63	30	18.93	30	19.78	60
Grade/CTBS/Q Test							
8:	Reading	36.57	30	35.60	30	36.08	60
	Language	38.07	30	38.93	30	38.50	60
	Arithmetic	42.83	30	51.57	30	47.20	60
	Study Skills	22.07	30	23.07	30	22.95	60
Grade/CTBS/Q Test							
10:	Reading	40.93	30	44.90	30	42.92	60
	Language	40.80	30	49.43	30	45.12	60
	Arithmetic	43.93	30	48.90	30	47.42	60
	Study Skills	25.03	30	24.60	30	24.82	60

These analyses were important because first year scores were collected as baseline data. Since only one significant difference was found, the researchers concluded that the treatment and comparison students were very similar in terms of achievement during the first program year before any substantial treatments were initiated.

The supplementary analysis carried out for the first year dealt with how the treatment and comparison students scored in relation to the national norms. As pointed out in the Second Interim Report, the students in both the treatment and comparison schools, for the most part, scored below the national norms (7). The only exception was that the kindergartners in both groups scored equal to or above norm on the TOBE subtests. Most of the first graders scored below norm on the TOBE and the second through eleventh graders continued to be below national norm on the CTBS/Q scores.

The number of students who scored below the national norms was expressed as a percentage of the total students in the tested grades in the treatment and comparison schools. Using that derived statistic confirmed the finding that most of the students above kindergarten scored below the norms in both groups, and further showed that the pattern of relationships to the norms was not consistent. Table 3 shows these percentages for grades K, 1, 5, 8, and 10.

Table 3: Percentage of Students Scoring Below National Norms: TOBE and CTBS/Q Subtests, 1972-1973*

Grade	Group	Language	Mathematics	Science	Social Studies
TOBE		%	%	%	%
K	Treatment	17.8	20.3	33.9	14.8
	Comparison	20.0	33.6	35.4	33.9
1	Treatment	72.1	75.2	72.2	78.4
	Comparison	68.1	75.5	72.0	83.3
CTBS/Q		Reading	Language	Arithmetic	Total Battery	Study Skills
		%	%	%	%	%
5	Treatment	95.9	90.6	87.2	93.9	94.8
	Comparison	96.7	88.4	92.9	94.2	96.8
8	Treatment	97.3	93.6	97.3	97.5	92.7
	Comparison	95.2	94.3	91.4	95.2	92.4
10	Treatment	90.0	93.3	93.4	93.3	88.4
	Comparison	83.8	85.1	91.9	90.5	89.0

*TOBE percentages are based on stanines, CTBS/Q percentages are based on grade equivalencies, thus direct relationships between the two cannot be ascertained from the percentages.

An examination of these percentages indicated that the pattern described in other studies was evident - students from this population deviated further and further from the norm as they progressed through school. For all grades tested, 83.2% of the treatment students, versus 77.3% of the comparison students, scored below the national norms that year.

Second Year Findings

The classroom means were computed for each subtest at each grade in all elementary schools, together with the mean scores on the subtests taken by the ninth and eleventh graders in both the treatment and comparison high schools. These means

are shown in Table 4 where schools A, B, C, D, and I were treatment schools, and E, F, and J were comparison schools.

Table 4: TOBE/CTBS/Q Raw Score Means by Grade and School: 1973-1974*

Grade/School		Language	Math	Science	Social Studies	
K:	A	18.30	19.02	17.44	17.37	
	B	21.63	20.47	19.33	21.03	
	C	16.59	17.82	15.91	16.44	
	D	20.85	19.24	18.19	18.21	
	E	19.55	17.66	15.79	16.77	
	F	19.50	18.87	17.84	16.99	
		Reading	Language	Math	Total Battery	Study Skills
1:	A	44.29	17.30	29.08	91.02	N.A.
	B	43.01	20.78	27.30	90.97	N.A.
	C	41.10	17.42	27.25	87.32	N.A.
	D	32.43	18.39	24.15	74.89	N.A.
	E	47.92	19.54	27.99	96.50	N.A.
	F	41.73	17.56	25.33	85.09	N.A.
2:	A	25.03	25.74	41.77	95.22	9.60
	B	21.03	21.89	30.69	73.65	8.43
	C	22.58	23.16	31.41	77.45	9.34
	D	19.34	16.46	29.07	64.79	7.58
	E	23.99	27.12	39.19	91.98	9.13
	F	23.34	25.21	32.41	79.52	8.41
3:	A	31.87	34.26	54.46	122.53	12.33
	B	30.59	33.20	45.96	109.87	11.12
	C	35.26	37.66	57.03	131.01	11.16
	D	30.91	34.87	50.66	119.33	10.90
	E	39.44	43.02	58.50	149.61	14.05
	F	29.91	35.40	45.81	111.55	11.97
4:	A	39.54	44.08	66.99	151.96	13.84
	B	39.21	43.59	62.31	149.04	15.02
	C	38.24	43.15	67.49	152.00	13.98
	D	34.01	37.62	57.54	131.11	12.48
	E	40.25	46.24	72.82	159.72	14.92
	F	35.78	39.82	62.17	137.80	13.25
5:	A	35.77	36.56	50.33	124.65	19.17
	B	34.05	35.13	49.41	118.61	18.55
	C	37.72	40.72	50.35	129.75	20.79
	D	33.97	33.86	46.75	117.84	18.74
	E	30.88	30.90	43.55	106.87	17.07
	F	38.61	39.51	51.47	127.05	20.55
6:	A	38.46	38.89	51.25	131.05	21.64
	B	39.62	40.68	59.35	143.56	24.31
	C	42.95	42.07	59.85	145.90	23.43
	D	41.59	42.35	53.61	139.02	20.57
	E	37.03	40.40	58.34	136.24	20.03
	F	39.69	44.41	56.17	141.04	24.32
9:	I	35.03	39.74	44.10	119.62	22.71
	J	31.92	34.40	37.25	105.11	19.14
11:	I	48.27	49.86	56.06	155.81	28.74
	J	42.39	45.72	44.50	133.32	24.41

*Data for grades K through 6 are mean classroom scores; data for grades 9 and 11 are mean scores for the entire grade. Schools A, B, C, D, and I were treatment schools; E, F, and J were comparison schools.

BEST COPY AVAILABLE

At the elementary level, a series of between school, one way ANOVAs were performed on the subtest means of the six schools at each grade. The results of these tests were presented in the Second Interim Report, which showed that significant differences between schools appeared on the language and social studies subtests of the TOBE at the kindergarten level, and on the language subtest of the CTBS/Q at grade five (7). A summary of the analyses of variance performed on these subtest data is presented in Appendix B1.

Although these findings demonstrated that significant differences existed in the classroom means among the six elementary schools, they did not indicate which school means were significantly higher than the others. Since different treatments were in effect in the four ESP schools during this year, it was necessary to isolate the significant differences to determine if any of these schools were involved.

These differences were located by applying Scheffe's post hoc multiple comparison technique to all combinations of school means for those subtests in which significant F ratios were found. In kindergarten, the significant differences existed only between two treatment schools. One of the schools (B) had the highest classroom mean on both the language and social studies subtests, the other school (C) had the lowest mean on both subtests among the six schools.

At the fifth grade, even though the ANOVA procedure revealed a significant difference in the classroom means among the six elementary schools on the language subtest, Scheffe's post hoc comparisons revealed no significant differences. Recognizing that the Scheffe method is very conservative, two somewhat less conservative tests were employed to locate the differences, namely, the Newman-Keuls and Tukey tests. Both of these revealed that a significant difference existed between the highest mean (40.72 for school C) and the lowest mean (30.90 for school E). The former was one of the treatment schools and the latter one of the comparison schools.

The mean scores of the high school students were analyzed in a 2 by 2 factorial design with the treatment group versus the comparison group as one dimension, and the ninth graders versus the eleventh graders as the other dimension. The results of the high school ANOVAs performed on the CTBS/Q subtests and total battery are presented in Appendix B2.

While the eleventh graders were expected to score significantly higher than the ninth graders, no differences were expected by group. The findings, however, revealed significant differences by group and by grade. The treatment students scored significantly higher than the comparison students on all subtests and the total battery, and the eleventh graders scored significantly higher than the ninth graders with one exception.

The performance of the treatment ninth graders on the mathematics subtest was especially noteworthy. It was pointed out in the findings for the previous year that the comparison eighth graders scored significantly higher than the ESP eighth graders on this subtest. Most of these eighth graders subsequently were promoted to the ninth grade in both the treatment and comparison high schools. Not only did the treatment ninth graders do significantly better than the comparison ninth graders that year, a reversal of the previous year, but the mean score of the treatment ninth graders (44.10) was nearly equal to the mean score of the comparison eleventh graders (44.50). Thus, within one school year, the treatment ninth graders made substantial gains in their mathematics skills as measured by the CTBS/Q.

In the supplementary analyses, an investigation of the changes in achievement scores from the first project year to the second was carried out for elementary and secondary treatment and comparison students. At the elementary level, changes in the second year mean scores of all treatment sixth graders, relative to their mean scores as fifth graders the year before, were calculated. The same analyses were carried out for the comparison sixth graders. Similar analyses were performed on the differences between the first and second year mean scores of the treatment eleventh graders (tenth the year before) against the mean score differences among the comparison students in the same grades. The mean scores analyzed together with the cell sizes are presented in Table 5.

An analysis of covariance technique (ANCOVA) was used to carry out the comparisons from year to year, employing the first year means as the pre test scores (covariate), and the second year means as the post test scores (dependent variable). The results of the ANCOVAs were presented in the Second Interim Report (7) which showed that significant differences were obtained only on the mathematics subtest at the elementary and secondary levels, and on the total battery at the secondary level. A summary of the analysis of covariance results for these subtests is repeated in Appendix B3.

At the elementary level, no significant differences were found on the mathematics subtest between the treatment and comparison fifth graders during the first year, nor between the treatment and comparison sixth graders in the second year. The ANCOVA results, however, showed that the mean scores of the treatment students from the first to the second year were significantly higher than those of their comparison counterparts. The treatment students gained 7.58 points over the first two years, which was significantly higher than the 4.95 points gained by the comparison students.

Table 5: Mean Raw Scores of Elementary and Secondary Students Analyzed from the First to the Second Project Year

CTBS/Q TESTS	Elementary					
	Treatment			Comparison		
	Year 1	Year 2		Year 1	Year 2	
	5th Grade $\bar{X}$	6th Grade $\bar{X}$	N	5th Grade $\bar{X}$	6th Grade $\bar{X}$	N
Mathematics	52.82	60.40	221	49.04	53.99	75
Language	38.58	43.92	212	38.99	43.37	76
Reading	31.05	41.72	219	37.76	41.29	76
Total Battery	130.30	149.44	203	125.58	139.38	71
Study Skills	20.06	22.47	222	19.62	21.21	78
	Secondary					
	10th Grade	11th Grade		10th Grade	11th Grade	
Mathematics	50.20	58.51	132	50.59	51.06	51
Language	45.09	47.14	145	49.06	46.80	49
Reading	44.12	44.89	145	47.00	44.80	49
Total Battery	140.56	154.98	130	147.06	144.43	49
Study Skills	25.67	22.84	135	25.05	21.47	59

NOTE: The means given are for all students who completed that test thus the total battery mean is not the sum of the means for the mathematics, language, and reading subtests.

At the secondary level, the findings from the first year revealed that although the mean score of the comparison tenth graders on the mathematics subtest was slightly higher than the mean score of the treatment tenth graders, the difference between the two groups was not significant. During the second year, however, the treatment eleventh graders scored significantly higher than comparison eleventh graders. This advance by the treatment students was also evident in the year to year ANCOVAs as treatment scored significantly higher than the comparison students from the first to the second year. The treatment students gained 8.31 points over this period, significantly higher than the .47 points gained by the comparison students.

Since no significant differences were found in the year to year changes on any subtests other than mathematics, it was posited that the high gain by the secondary treatment students on this subtest contributed the major portion of the increase in their total battery mean score. The experimental change was significantly higher than the year to year change realized by the comparison students.

Third Year Findings

In accordance with the research design, the achievement data collected during the third year were analyzed using both the first and second research questions pertaining to group and sex.

Both questions were analyzed together in a series of 2 by 2 ANOVAs performed on the mean scores of the students in each grade, K through 12. As in the previous years, all subtests of the TOBE or CTBS/Q were examined within each grade, and if significant differences occurred from grades K through 6 in favor of the treatment students, the data were further analyzed to find out which school means were involved. A supplementary analysis for the third project year dealt with an investigation of student performance by quartile.

The pattern of findings from the 2 by 2 ANOVAs performed at grades K through 6 and 7 through 12 are presented in Tables 6 and 7, respectively. These tables show where significant differences at the .05 level were found (signified by an S) on the various subtests in each grade at the elementary and secondary levels. For interpretative purposes, the tables also indicate the direction of the significant differences in terms of the treatment and comparison groups, males and females, as well as whether or not the interaction effects between group and sex factors were significant. For all significant results found at each grade in these tables, a summary of the analysis of variance by group and sex, together with the appropriate means, is presented in Appendices C1 through C13.

The data showed that the treatment students scored significantly higher than the comparison students on all subtests at grades K, 1, 10, 11, and 12, and on some subtests at grades 8 and 9. The comparison students scored significantly higher on scattered subtests at grades 3, 4, and 5. In regard to the findings by sex, the males scored significantly higher than the females only on the science subtest in kindergarten. In the twenty-three other cases where sex was significant, the females scored higher than the males. These differences became prevalent at grades 6 and 7, peaked at the 8th grade, and continued to exist through the eleventh grade. Also, the females were predominantly higher on the language subtest. All of the significant group and sex interaction effects occurred at grades 7 and 8.

Since the ESP students obtained significantly higher means on the TOBE than did the comparison students at grades K and 1, further analyses were carried out to determine more precisely where the significant differences occurred among the six elementary schools. The data for these grades were rerun in a one way ANOVA and Scheffe post hoc multiple comparisons were performed on all pairs of means among the six schools on each subtest. In kindergarten, the analysis revealed that three of the treatment schools were significantly higher than one of the comparison schools on the mathematics subtest; two treatment schools were higher than one comparison school on the social studies subtest; and one treatment school was higher than one comparison school on the language subtest. The pattern was similar at the first grade.

Table 6: Summary and Direction of Findings by Group and Sex: Grades K-6, 1974-1975, 1975-1976

GRADE/TEST*	1974-1975					1975-1976				
	GROUP		SEX		INTER-	GROUP		SEX		INTER-
	ESP>C	C>ESP	M>F	F>M	ACTION	ESP>C	C>ESP	M>F	F>M	ACTION
K										
Language	S		-	-	-	-	-		S	-
Mathematics	S		-	-	-	S		-	-	-
Science	S		S		-	NA	NA	NA	NA	NA
Soc. Studies	S		-	-	-	NA	NA	NA	NA	NA
1										
Language	S			S	-	-	-	-	-	-
Mathematics	S		-	-	-	S		-	-	-
Science	S		-	-	-	NA	NA	NA	NA	NA
Soc. Studies	S		-	-	-	NA	NA	NA	NA	NA
2										
Reading	-	-	-	-	-		S	-	-	-
Language	-	-	-	-	-		S		S	-
Arithmetic	-	-	-	-	-		S	-	-	-
TotBattery	-	-	-	-	-		S		S	-
Study Skills	-	-	-	-	-		S	-	-	-
3										
Reading	-	-	-	-	-	-	-	-	-	-
Language	-	-	-	-	-	-	-		S	-
Arithmetic	-	-	-	-	-	-	-	-	-	-
TotBattery	-	-	-	-	-		-	-	-	-
Study Skills		S	-	-	-	-	-	-	-	-
4										
Reading	-	-	-	-	-	-	-	-	-	-
Language	-	-		S	-		S	-	-	-
Arithmetic	-	-	-	-	-		S	-	-	-
TotBattery	-	-	-	-	-	-	-	-	-	-
Study Skills		S	-	-	-	-	-	-	-	-
5										
Reading	-	-	-	-	-	-	-	-	-	-
Language		S		S	-		S	-	-	-
Arithmetic	-	-	-	-	-	-	-	S		S
TotBattery		S	-	-	-		S	-	-	-
Study Skills		S	-	-	-	-	-	-	-	-
6										
Reading	-	-	-	-	-		S	-	-	-
Language	-	-		S	-		S	-	-	-
Arithmetic	-	-		S	-		S	-	-	-
TotBattery	-	-		S	-		S	-	-	-
Study Skills	-	-	-	-	-		S	-	-	-

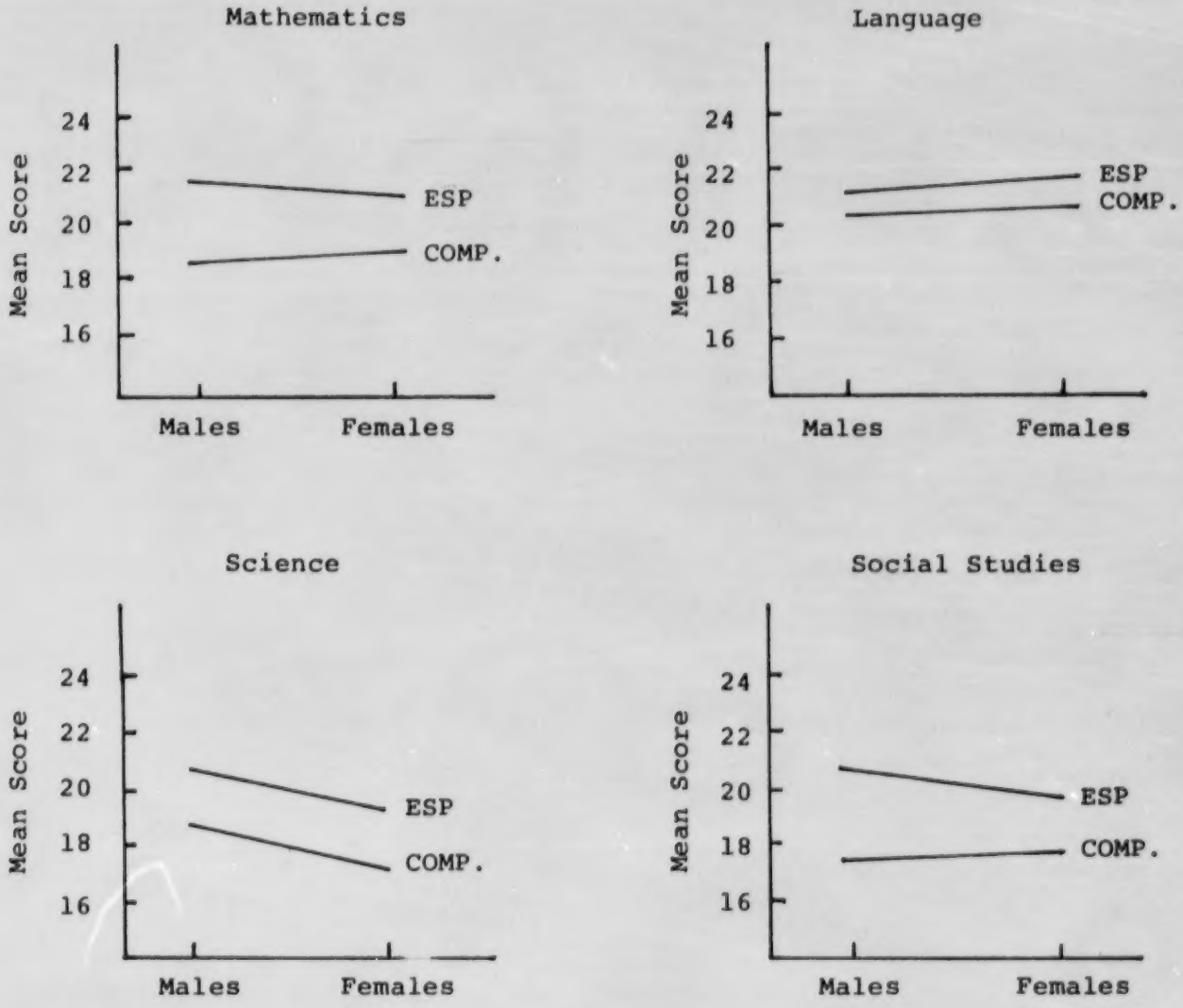
*For grades K and 1, the subtests apply to the TOBE; for grades 2 through 6, the subtests apply to the CTBS/Q.
NA=Not applicable; comparison K and 1 tested only on TOBE language and mathematics subtests in 1975-1976.

Table 7: Summary and Direction of Findings by Group and Sex: Grades 7-12, 1974-1975, 1975-1976

GRADE/TEST*	1974-1975					1975-1976				
	GROUP		SEX		INTER-ACTION	GROUP		SEX		INTER-ACTION
	ESP>C	C>ESP	M>F	F>M		ESP>C	C>ESP	M>F	F>M	
7										
Reading	-	-	-	-	-	-	-	-	-	-
Language	-	-		S	S	-	-		S	-
Arithmetic	-	-	-	-	-	-	-	-	-	-
TotBat	-	-		S	-	-	-	-	-	-
Study Skills	-	-		S	S	-	-	-	-	-
8										
Reading	S		-	-	-	-	-	-	-	-
Language	-	-		S	S	-	-		S	-
Arithmetic	S			S	-		S	-	-	-
TotBattery	S			S	S	-	-		S	-
Study Skills	S			S	S		S		S	-
9										
Reading	S			S	-	-	-	-	-	-
Language	-	-		S	-		S		S	-
Arithmetic	S		-	-	-	-	-	-	-	-
TotBattery	-	-		S	-	-	-		S	-
Study Skills	-	-	-	-	-	-	-	-	-	-
10										
Reading	S			S	-	S		-	-	-
Language	S			S	-	-	-		S	-
Arithmetic	S		-	-	-	-	-	-	-	-
TotBattery	S			S	-	-	-		S	-
Study Skills	S		-	-	-	S		-	-	-
11										
Reading	S			S	-	S		-	-	-
Language	S			S	-	-	-		S	-
Arithmetic	S		-	-	-	S		-	-	-
TotBattery	S			S	-	-	-	-	-	-
Study Skills	S		-	-	-	-	-	-	-	-
12										
Reading	S		-	-	-	S		-	-	-
Language	S			S	-	-	-		S	-
Arithmetic	S		-	-	-	-	-	-	-	-
TotBattery	S		-	-	-	-	-	-	-	-
Study Skills	S		-	-	-	S		-	-	-

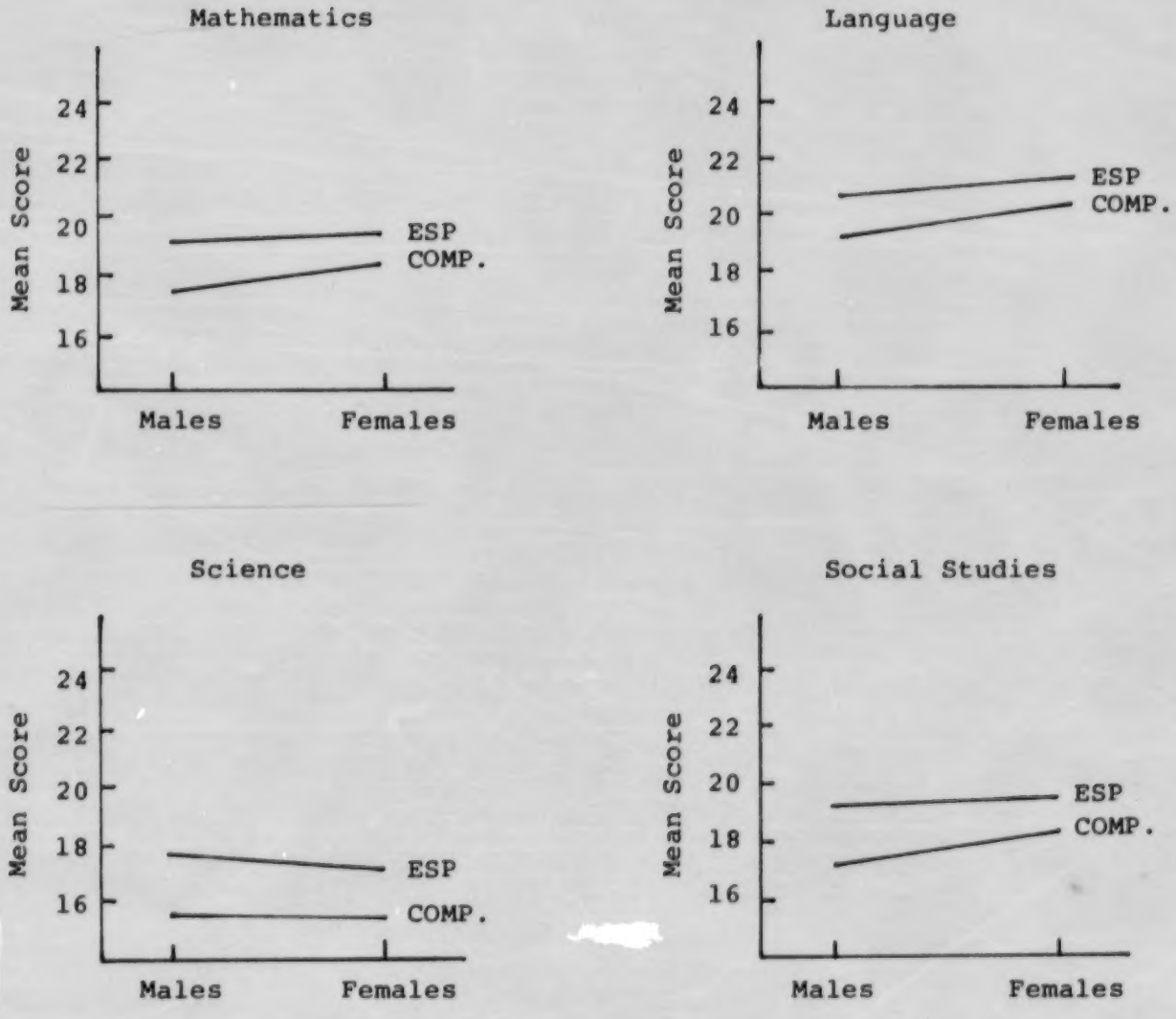
*For grades 7 through 12, the subtests apply to the CTBS/Q.

Figure 1: Kindergarten TOBE Means by Group and Sex: 1974-1975



Figures 1 and 2 graphically illustrate the positions of the class means of the kindergarten and first grade students, respectively, by group and sex for all subtests of the TOBE. The widest difference between ESP and comparison groups existed on the social studies subtest at kindergarten. The ESP males scored 3.3 points higher than the comparison males. The average difference between all ESP and comparison groups on

Figure 2: First Grade TOBE Means by Group and Sex: 1974-1975



all four subtests at these two grades was 1.9 points. Therefore, even though the differences between the two groups were statistically significant in favor of the ESP students, the difference did not appear to have any practical implications for ESP program restructuring.*

The analyses of the CTBS/Q results from grades two through twelve revealed that some of these differences probably would have practical significant. Table 8 sets forth the absolute differences in the mean scores of the subtests at all grades where statistical significance was indicated in Tables 6 and 7.

Table 8: Absolute Differences between Statistically Significant Treatment and Comparison Raw Score Means, CTBS/Q: 1974-1975

Grade	Subtest	Difference in Means	Higher Group
3	Study Skills	1.61	Comparison
4	Study Skills	1.87	Comparison
5	Total Battery	10.99	Comparison
	Language	5.57	Comparison
	Study Skills	2.30	Comparison
8	Total Battery	8.95	ESP
	Reading	3.93	ESP
	Arithmetic	3.44	ESP
	Study Skills	2.17	ESP
9	Reading	2.25	ESP
	Arithmetic	2.28	ESP
10	Total Battery	11.68	ESP
	Reading	4.18	ESP
	Language	4.69	ESP
	Arithmetic	2.77	ESP
	Study Skills	2.75	ESP
11	Total Battery	17.78	ESP
	Reading	5.94	ESP
	Language	4.74	ESP
	Arithmetic	7.10	ESP
	Study Skills	4.73	ESP
12	Total Battery	11.98	ESP
	Reading	3.64	ESP
	Language	2.90	ESP
	Arithmetic	5.43	ESP
	Study Skills	2.28	ESP

NOTE: See Appendices C1 - C13 for additional data on these means and cell sizes.

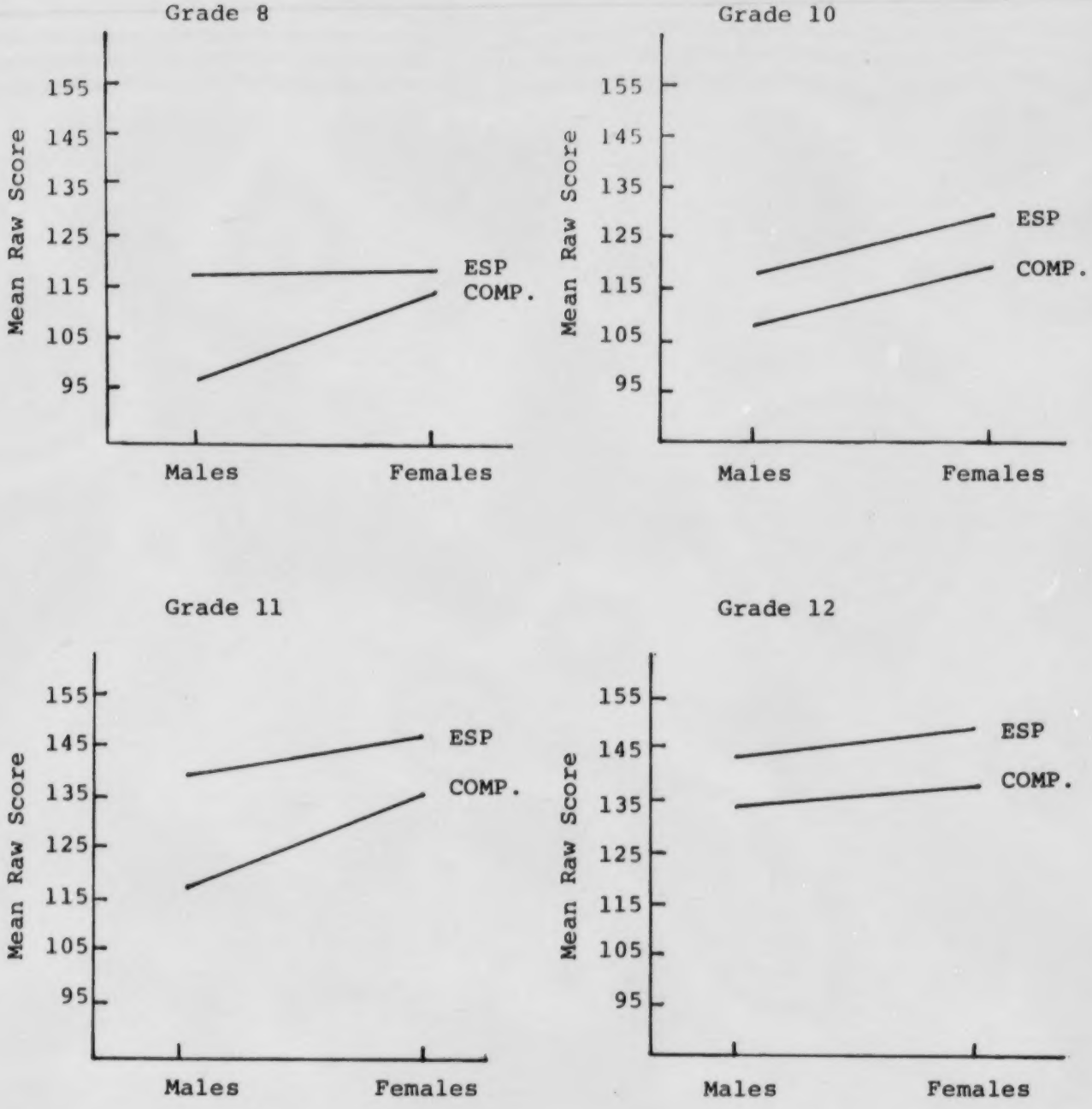
*At all but the highest levels, a two point difference on the CTBS/Q total battery score would generally fall within a tenth of a grade equivalency (GE), e.g., total battery scores of 193, 194, and 195 are encompassed in a GE of 5.8 and 196, 197, and 198 are in a GE of 5.9. Some tenths of a GE range, representing roughly a month of instruction, include up to five points on the total battery score. (10)

Using a difference in the mean scores of 5 (roughly the value of a month of instruction) or more points as having practical significance, those between group differences on the total battery at grades 5, 8, 10, 11, and 12 would appear to be important ones. The comparison group was higher than the ESP group at grade 5, and the ESP group was higher than the comparison group at grades 8, 10, 11, and 12 for the third project year.

The differences in the total battery means for grades 8-12, by group and sex, are shown graphically in Figure 3. As the graphs demonstrate, the treatment students performed substantially better than the comparison students in these cases except for the ESP females, who scored only slightly better than the comparison females at the eighth grade. The minimum difference in mean scores between any group of ESP and comparison students at these grades was 9.44 points, except between ESP and comparison females at the eighth grade where the difference was 2.32 points.

It is also evident from the graphs that the females, whether in the ESP or the comparison group, generally scored higher than the males. The females in both groups scored much higher on the language subtest, thereby increasing the amount of their total battery scores over the males.

Figure 3: Mean Raw Scores on CTBS/Q Total Battery by Group and Sex, Grades 8, 10, 11, and 12: 1974-1975



To determine whether the group or sex effect was stronger at these grades, the eta squared coefficient was derived in conjunction with the 2 by 2 ANOVAs performed on the total battery scores. This coefficient has gained widespread acceptance among educational researchers as a postmortem, descriptive measure of the strength of association between effects and dependent variables. (28). Table 9 shows the eta squared coefficient for the group and sex effects at the four grades.

Table 9: Eta Squared Coefficients on CTBS/Q Total Battery, Grades 8, 10, 11, and 12: 1974-1975

Grade	Effect	Eta^2
8	Group	.014
	Sex	.014
10	Group	.022
	Sex	.020
11	Group	.048
	Sex	.036
12	Group	.020
	Sex	.004

Although no eta squared was very strong, the group effect was greater than the sex effect at all grades except the eighth. Indeed, a significant interaction effect occurred at the eighth grade. As the graph in Figure 3 suggests, the overall difference in scores at the eighth grade was partly due to the poor performance of the comparison males rather than just to the superior performance of all the ESP students. At grades 10, 11, and 12, however, both the male and female ESP students scored higher than their counterparts in the comparison schools.

Probing beyond the total battery scores at these grades revealed that the next highest differences in mean scores between the treatment and comparison students, when sex was disregarded, were almost always found on the reading and/or arithmetic subtests. Looking back at the absolute differences in raw score means in Table 8, the highest differences in favor of the ESP students occurred across the three subtests used to calculate the total battery score: reading, language, and arithmetic. Several of these differences were over five points, indicating some probable practical as well as statistical significance.

In summary, the results of the analyses of the data by group and sex for the third year revealed that although the treatment students scored significantly higher than the comparison students on many of the subtests at grades K and 1, and 8 through 12, the statistically significant differences did not always indicate a practical significance as well. Differences in mean raw scores of more than five points in favor of the ESP students were found only at grades 10 through 12. Nearly all of these cases involved the reading and arithmetic subtests. The comparison group was significantly higher than the ESP group at grades 3, 4, and 5 on 5 of the 15 subtests. Scores on the study skills subtest were the only ones significantly higher at the third and fourth grades. In the fifth grade, study skills, language, and total battery scores were higher for the comparison group also.

In the supplementary analysis carried out during the third project year, the researchers attempted to determine if the project treatments had affected the students by performance groups. Performance group was defined in terms of the distribution of scores by quartile.* As with the analysis related to sex, the ESP project had not focused its treatment on either the students of higher or lower performance independently, but one group could have been more affected by the methods and curricular content.

Grades K, 1, 8, and 10 were selected for these analyses since the students in these grades received most of the treatments during the third year, and additionally, the elementary, middle, and secondary levels were represented. At each grade, t tests were performed on the mean scores of all TOBE or CTBS/Q subtests for each performance group between the treatment and comparison students. The pattern of results was then analyzed across the four grades to determine if the treatment students did consistently better than the comparison students in any of the performance quartiles. The results are shown in Table 10.

As the figures in the table demonstrate, there was no consistent pattern of results by subtest or quartile. Of the nine significant differences found, the ESP students scored higher than the comparison students on the social studies subtest at grade 1, and on the study skills subtest at grades 8 and 10. By quartile, five of the nine significant differences were found at the fourth quartile, two at the first quartile, and one each at the second and third quartiles. Such results are not strong enough to conclude that the program affected the higher or lower performance students differently.

*Since whole number scores were used, the distribution was not exactly evenly divided among the quartiles; the variation was slight, however.

Table 10: Summary of Significant Differences between ESP and Comparison Mean Scores by Quartile, TOBE-CTBS/Q: 1974-1975

Grade	Subtest	Quartile	Treatment $\bar{X}$	Treatment sd	Comparison $\bar{X}$	Comparison sd	p
K	Science	4	25.07	1.39	24.27	1.19	.01
1	Social Studies	1	12.61	2.57	11.13	3.04	.05
	Social Studies	4	24.89	1.56	24.21	1.25	.05
8	Language	3	41.09	2.79	39.67	2.90	.01
	Arithmetic	1	25.22	3.77	26.64	3.95	.05
	Total Battery	4	171.34	24.22	160.10	25.76	.05
	Study Skills	4	32.20	4.99	30.38	4.76	.05
10	Study Skills	2	18.28	1.66	17.39	1.68	.01
	Study Skills	4	33.39	3.98	31.55	3.65	.01

Fourth Year Findings

Virtually the same amount of achievement data was collected the fourth year as during the previous year. The only exception to this occurred in the kindergarten and first grades in the two comparison elementary schools. By decision of the district, the comparison students in these grades were tested only with the mathematics and language subtests of the TOBE and not the science and social studies subtests to conform with other testing activities being conducted.

As in the previous year, 2 by 2 ANOVAs were used to determine if group and/or sex differences in achievement test scores existed at each grade, K through 12. The supplementary analyses carried out in the first year were repeated the fourth year to determine if the percentage of students scoring below national norms had changed over the four years.

Fewer significant differences in favor of the ESP students were found the fourth year. Summary tables of analyses of variance by group and sex, together with the appropriate means, are presented in Appendices D1 through D13. In the third project year, twenty-five significant differences were found in favor of the ESP students (excluding those in science and social studies at grades K and 1), whereas only eight occurred in the fourth year. (Refer again to Tables 6 and 7.) These eight differences existed either in the primary or the secondary grades: on the mathematics subtest at grades K and 1, and on the reading, arithmetic, or study skills subtests at grades 10 through 12. The comparison students, on the other hand, scored significantly higher on 19 of the tests; these included most of the achievement areas, including the total battery, at grades 2, 4, 5, and 6; on the arithmetic and study skills subtests at grade 8; and on the language subtest at grade 9.

In regard to the fourth year findings by sex, the males scored significantly higher than the females only on the arithmetic subtest at the fifth grade. In the fourteen other cases where sex was significant, the females again scored higher than the males. As in the previous year, these differences were more prevalent at the eighth grade and nearly always on the language subtest. The only significant interaction effect on any of the tests performed occurred on the arithmetic subtest at the fifth grade.

Since the treatment students obtained a significantly higher mathematics score at grades K and 1, the data for these grades were rerun in a one way ANOVA with post hoc multiple comparisons using the Scheffe procedure to determine where the differences lay among the six elementary schools. For grade K, the analysis of TOBE scores revealed that the means of two of the treatment schools were higher than the mean of one of the comparison schools. The same schools were involved as in the previous year when three treatment schools were higher than the comparison school. At the first grade, three of the four treatment school means were higher than one comparison school mean. This was in contrast to one treatment school being higher than one comparison school the year before.

Despite these differences in individual schools, however, the overall difference between the treatment and comparison students on the mathematics subtest was less than two points at both the kindergarten and first grade. (See Appendices D1 and D2.) This situation indicated that once again, although there was a statistically significant difference in favor of the treatment students, it may have been of little practical value. Thus, the pattern of results over the last three years of the program appeared to be one where grade K in one treatment school showed differences of both statistical and practical significance for the students in that school, but for kindergarten in the other treatment schools and for first grade in all schools, differences of statistical rather than practical value occurred.

Reiterating the trend of data mentioned for the other years, the comparison students scored significantly higher than the treatment students in most of the achievement areas at grades 2, 4, 5, and 6. Post hoc comparisons carried out on the relevant means of all the elementary schools at these grades revealed that the students in one of the comparison schools scored significantly higher than those in any of the other schools, which accounted for the overall pattern of the superior comparison group performance.

At the secondary level, the findings showed significant differences between the mean scores of the treatment and comparison students only on the reading, arithmetic, or study skills subtests at grades ten through twelve. And only a few of these differences were greater than five points, suggesting

that they were of little practical utility (Appendices D11 to D13). This pattern represented a change from the previous year when the ESP students were significantly higher than the comparison students, both in the statistical and practical sense, in all areas of achievement, including the total battery.

Table 11: Percentage of Students Scoring Below National Norms: TOBE and CTBS/Q Subtests, 1975-1976*

Grade		Language	Mathematics	Science	Social St.
TOBE	Group	%	%	%	%
K	Treatment	22.7	18.7	32.3	19.3
	Comparison	36.0	21.4	**	**
1	Treatment	59.0	50.0	52.5	65.1
	Comparison	71.1	54.4	**	**

CTBS/Q		Reading	Language	Arithmetic	Total Battery	Study Skills
		%	%	%	%	%
5	Treatment	89.3	86.1	88.7	89.9	87.7
	Comparison	89.7	80.4	85.3	85.9	87.6
8	Treatment	95.8	94.4	97.1	97.4	92.8
	Comparison	98.3	96.1	97.5	98.8	92.8
10	Treatment	89.7	90.9	95.0	95.7	83.7
	Comparison	95.5	95.5	96.1	97.8	93.9

*TOBE percentages were based on stanines; CTBS/Q percentages were based on the grade equivalencies, and thus are not directly comparable with TOBE stanine data.
**Not applicable.

In the supplementary analyses for this year, the investigation of the percentage of students scoring below national norms carried out the first year were repeated to determine if the relative position of the treatment and comparison groups had changed over the four years. The percentage of students scoring below national norms the first year in grades K, 1, 5, 8, and 10 was presented in Table 2. Comparing the percentages of treatment students below national norms between the first and fourth years, little change was noted. A tendency toward improvement in mathematics was found in grades K and 1. The percentages of comparison students below national norms did not present a consistent pattern across all grades studies. Grades 8 and 10, however, appeared to have worsened and especially so at grade 10.

Viewing the means in Appendices D1, D2, D6, D9, D11, the treatment and comparison students are quite similar in overall pattern. While trend lines were not calculated for these groups, the greater percentage below norm, plus the relatively stable means suggested that some bimodality in the grade 8 and 10 score patterns occurred for the comparison group. That is to say, some students improved their scores enough to offset the larger number that did not improve.

Summarizing the results of this supplemental analysis, little change in the position of the treatment group occurred. Their relationship to national norms varied little over the four years. The comparison group's position with regard to the norms was also about the same in the elementary grades analyzed but a trend toward a greater number below norm in secondary was noted. Overall, the Edgewood scores changed little; a progressively widening gap occurred through the grades.

SUMMARY AND DISCUSSION

Analyses of the achievement scores of sample treatment and comparison students were performed each of the four years of the experimental program to determine if the presence of treatment programs was associated with changes in mean scores. Since the treatments were not of uniform emphasis across the four years, special attention was paid to those grades in which an ESP program was in existence.

First Project Year

The first year, 1972-1973, was considered the baseline year. Only a few ESP treatments were begun because of the staff's concentration on program organization. Most of the treatments were short term, a semester or less, and were not expected to have meaningfully affected the achievement scores.

This was borne out when no important differences between the treatment and comparison achievement score means were found when the spring test scores were analyzed. One subtest, arithmetic, at one grade, was significantly higher for the comparison group. Treatment and comparison students were thus judged to be essentially alike in their levels of achievement.

Second Project Year

Some treatments were instituted the second year but most of these were again short term and were not uniform across the grades. A special kindergarten program, addressing all the subjects, was instituted at one ESP school. Diagnostic reading was instituted in some treatment schools at the elementary

level. Best paced mathematics was begun with a number of classrooms in both the middle and secondary ESP schools. All other treatments appeared to be exploratory and transitory in application.

The treatment children scored significantly higher than those in the comparison schools on all the TOBE subtests. The ESP fifth graders' scores were higher in language achievement than were the comparison students' scores. The secondary students tested in the ESP schools scored higher on all the CTBS/Q subtests and total battery than did the comparison students but mathematics apparently had contributed most of the total battery increase. The gains in mathematics from the first to the second year by ESP students were greater than those obtained by the comparison students. The best paced mathematics program was implemented in all secondary mathematics classes the second year.

At some grades and on some subtests, however, the comparison students scored lower than they had the year previously. Furthermore, the schools within each group did not score uniformly in either direction. The presence of the kindergarten program in one particular school substantially raised the treatment group's means and also appeared to be associated with the increase of treatment group levels over those of comparison students. The existence of a diagnostic reading program, while not uniformly associated with augmented reading and language subtest scores, seemed to be related to substantial increases in two ESP schools. Several kinds of short term and one full year treatment program, best paced mathematics, were available in the ESP high school. Treatment scores were significantly higher there than the comparison scores in all the subtests and the total battery but the higher mathematics scores appeared primarily responsible for the treatment-comparison differences in the total battery scores. A relationship between the best paced mathematics treatment and the increased scores seemed to be tenable.

Third Project Year

The kindergarten program offered at the one school was continued even though at a lesser intensity than previously conducted. ESP announced concentration on that grade, plus the first and second grades, for its treatments during the third year. A part of the middle school was organized into integrated subject matter work and the students in that special program were taught by teams of teachers working to produce materials for them and to improve the general interaction between the teachers and students. The best paced mathematics program was carried out in all the mathematics classrooms in the middle and secondary schools. All other treatments were short term, sometimes lasting only six weeks, or were sporadic in grade and school application.

The ESP kindergarten and first grade continued to be significantly higher in achievement than were comparison students in the same grades. Such was not the case in grades 2 through 7. The only significant differences at these grades were in favor of the comparison students and most of the differences occurred on the study skills subtest. Thus the generalized treatment of inservice for teachers, more assistance at the treatment schools, and the provision of some materials, was not associated with measurable score increases.

Significant differences in favor of the ESP students did occur at grades 8 through 12 on most of the achievement subtest and total battery scores. The research controlled for sex the third year and many of the significantly different scores between the two groups were also accompanied by significantly higher scores for females, especially those in ESP, who were substantially highest. An examination of the orientation treatment programs, however, showed no discernable emphases in the instruction toward the females. Similarly, the comparison males often scored substantially lower than the other groups and, again, no instructional differences could be identified which could account for this difference.

An exception to this trend was mathematics, in which the comparison boys frequently scored higher than girls. A previous study had shown the boys to be absent more frequently than girls; however, that same research found no important difference in the treatment and comparison absenteeism rates.

Fourth Project Year

The ESP treatments were considerably expanded the fourth year, 1975-1976. The mathematics program, a revised version of the best paced program into a spiral arrangement, was generally in effect across the grades. Individualized instruction, sometimes through student contracts, was instituted in the language arts program. The middle school treatment program, revised extensively from the previous year, was expanded to include all the students in the ESP school. Some additional treatments were added to the secondary program and all ninth and tenth, plus all students in mathematics, were receiving some ESP treatments. A confounding change happened in the comparison high school; a large majority of the students in that school were put into a federally funded, individualized instruction program in science, social studies, reading, English, and mathematics. The materials for all the subjects except mathematics were teacher-consultant produced; mathematics was taught through a commercial program.

In kindergarten and first grade, the treatment students were significantly higher in mathematics than were comparison students but not in language. At second grade, the comparison students were higher than treatment students on all the subtests

and the total battery. There were no differences at third grade between the two groups. There were several statistically different scores at fourth through sixth grades, all in favor of the comparison students. Seventh grade data showed no differences by group. Isolated differences in favor of comparison occurred at the eighth and ninth grades; scattered differences in favor of the treatment students were found in the remaining grades. Although they were fewer than in the previous year, a number of sex differences, all in favor of females, remained in 1975-1976. It should also be pointed out that the treatment students, in many cases throughout the entire K-12 span, scored lower in the fourth year than each grade had scored the year before on some subtests and often on the total battery. Too, the several elementary schools still showed significantly different means, with one comparison school still highest and some treatment schools decidedly lower. An examination of the programs in these schools did not provide a consistent explanation for these differences.

Many of the data patterns established in previous years were interrupted in 1975-1976. One factor which may have affected achievement levels was the infusion, involving new materials and methods of relatively massive treatments, into most of the ESP classrooms. Some of the statistical differences occurred, or were not repeated from the previous year, because the ESP students frequently scored lower than they had before. That is to say, some statistical changes happened because of these lowered ESP scores, not because the comparison students had raised their scores substantially. The massive treatment program may have been responsible for a portion of the lowered scores. That is, learning to teach the new materials took time; adjusting to contracting, for example, has often been shown to cause students difficulty. It is possible, then, that some of the statistical changes were related to the amount and intensity of the new ESP treatment introductions.

A second factor that could have affected achievement was the conversion of the traditional approach in the comparison secondary school to individualized instruction. The fewer significant differences at the secondary level, with most but not all of them in favor of the treatment students, may have been associated with this change. That is, it would seem possible that increasing the amount of individualized instruction in the comparison school might have accounted for the reduced number of differences in favor of the treatment school. However, the statistically different subtest scores were not uniform; thus, attributing effect from this instructional change would have to be tenuous at best.

Recapitulation

A series of cross sectional analyses were performed on the subtest and total battery mean scores, when applicable, from

samples of treatment and comparison students at the grades where the maximum of treatment effect might be encountered. The samples were carefully selected and generally they were like the larger groups from which they were taken; only minor differences were found.

Looking at the results across the four years, even though the present study did not attempt analyses of the longitudinal data*, could be summarized, by variable, as:

1. The four treatment elementary and the two comparison elementary schools did not perform uniformly; especially important to this study, the ESP treatment programs did not appear to have uniformly affected the achievement scores of the youngsters in the different ESP schools.
2. The treatment and comparison students, when viewed as groups, did not present uniform patterns of differences; kindergarten scores and secondary mathematics scores tended to be higher for the treatment group but neither of these was consistent.
3. The achievement means of the treatment and comparison groups did not remain stable across the four years; the comparison group showed some losses the third year, the treatment groups showed some losses the fourth year.
4. A relatively large number of sex differences was encountered in the analyses for the last two years; these were almost entirely due to females, and especially the ESP females, scoring higher than the males and higher than either sex in the comparison group. ESP males, however, generally scored about the same as the comparison females and generally higher than the comparison males.
5. When examined by quartiles, the results were scattered and inconsistent; gains by any one quartile, either in treatment or comparison, showed no convincing pattern of differences.
6. When selected grades were examined to determine the percentage of students below the national norms, the treatment group appeared to have a smaller proportion below norm at the end of the four years than had been there at the beginning of the program and, additionally, a smaller proportion of the secondary students was below norm than in the comparison group.
7. Neither the treatment nor comparison group changed noticeably from the pattern of becoming increasingly further behind the national norm as the students progressed through grades K-12.

*The longitudinal and some additional cross sectional analyses were reported in Cervantes & Bernal (6).

When these mostly inconsistent patterns are taken into account, the conclusion must be that the experimental program was unable to effect general academic achievement increases from its treatments.

Some questions remained unanswered, chief among these being why the female pupils attending ESP schools so consistently achieved higher than the others. Examinations of the curricula and methods showed no overt inclinations toward the females nor the inclusion of unusual content that would detrimentally affect the learning of males. The comparison group differences were also unexplained. The possibility remained, then, that some findings from the general literature, which suggest that girls are better able to cope with innovation and/or that girls' learning styles are more suited to individualization, may have been involved in the present case, as well as some findings that Mexican American girls generally score higher than boys.

Another perplexing problem in the research was the third year drop in achievement levels for the comparison students and the fourth year drop for those in the treatment schools. While by no means absolute, nor across every subtest, both were consistent enough to demand examination. No important differences due to sampling could be ascertained. Mobility rates appeared to have remained about the same through the years and thus did not offer an easy explanation.

Although completely speculative, an obvious question is whether the results would have been more uniform had the treatments been more consistently applied. The two years of the kindergarten and best paced mathematics, accompanied by generally increased scores, would tend toward an affirmative answer. A three or four year application, which did not occur in either case, might have produced salutory results. This must remain only as a possibility but worthy of examination in future experimental programs.

In brief, this experimental program, during its first four years of funding, introduced a variety of treatments over different grades and for varying lengths of time. Some consistency of treatment was observed in the general approach and in the staff development training supplied to teachers. When viewed annually, some significantly higher ESP achievement scores appeared to be associated with treatments in the particular subjects and grades where the differences between ESP and the comparison groups were noted. ESP, however, tended to change its treatments often or at least to modify them considerably, and thus persistence of treatment effect was not encountered. In the aggregate, the considerable inconsistencies led to the general conclusion that the experimental program had not substantially improved the achievement scores of its pupils when viewed by particular years of program operation (i.e., cumulative impact), subject matter, grade, or quartile ranges, in relation to the Edgewood comparison students that did not participate in the program.

BIBLIOGRAPHY

1. ANDERSON, JAMES G. & WILLIAMS, M. Socio-cultural determinants of achievement among Mexican American students. Las Cruces, Educational Resources Information Center, New Mexico State University, 1968.

2. BROUSSARD, VERNON. The effect of an individualized approach on the academic achievement in mathematics of inner city children. Doctoral dissertation. East Lansing, Michigan State University, 1971.

3. CAMPBELL, DONALD T. & STANLEY, JULIAN C. Experimental and quasi-experimental designs for research. Chicago, Rand McNally, 1966.

4. CARTER, THOMAS P. The negative self-concept of Mexican American students. School and Society (96): 217-219. March 30, 1968.

5. CERVANTES, ROBERT A. Self-concept, locus of control, and achievement in Mexican American pupils. Doctoral dissertation. Union Graduate School - West, San Francisco, California, October 1975.

6. _____ & BERNAL, HELEN H. Psychosocial growth and academic achievement in Mexican American students. San Antonio, Texas, Development Associates. 1976.

7. CERVANTES, ROBERT A. & JONES, EARL (eds.) Second interim report on the Edgewood School Plan. San Antonio, Texas, Development Associates, January 1976.

8. CERVANTES, ROBERT A. & WELCH, RICHARD O. Self-concept and motivation of Chicano students. Paper presented at meeting of the Rocky Mountain Educational Research Association. Albuquerque, New Mexico, November 1, 1974.

9. COLEMAN, J. S. ET AL. Equality of educational opportunity. Washington, DC, Superintendent of Public Documents, 1966.

10. CTB/MCGRAW-HILL. Comprehensive Tests of Basic Skills, form Q, all levels: technical bulletin. Monterey, California, 1968.

11. DARCY, N. T. Bilingualism and measurement of intelligence: review of a decade of research. Journal of Genetic Psychology (103):259-282. December 1963.

12. DEL BUONO, X. ANTONIO. The relationship of bilingual/bicultural instruction to achievement and self-concept of seventh grade Mexican American students. Doctoral dissertation. East Lansing, Michigan State University, 1971.

13. DEVELOPMENT ASSOCIATES. First interim report on the Edgewood School Plan. San Antonio, Texas, 1972.

14. EDGEWOOD INDEPENDENT SCHOOL DISTRICT. Final plan: experimental schools project. San Antonio, Texas, May 1972.

15. _____. Personalizing instruction models and program emphasis. San Antonio, Texas, Edgewood School Plan, Spring 1975.

16. _____. Philosophy of communication skills. San Antonio, Texas, Edgewood School Plan, 1974.

17. FENNEMA, ELIZABETH & SHERMAN, JULIA. Sex-related differences in mathematics achievement, spatial visualization and affective factors. American Educational Research Journal 14(1):51-71. Winter 1977.

18. GEZI, KAL. Bilingual-bicultural education: a review of relevant research. California Journal of Educational Research XXV(5):223-239. 1974.

19. GREBLER, LEO. The schooling gap: signs of progress. Los Angeles, Division of Research, Graduate School of Business, University of California, 1967. (Mexican American Study Project Advance Report 7)

20. _____; MOORE, JOAN W. & GUZMAN, RALPH C. The Mexican American people. New York, The Free Press, 1970.

21. HARNISCHFEGER, ANNEGRET & WILEY, DAVID. Achievement scores drop. So what? Educational Researcher V(3). March 1976.

22. JONES, EARL. A study of curriculum and instruction in the Edgewood School Plan. San Antonio, Texas, Development Associates. (in preparation)

23. JONES, EARL. Evaluation and assessment report: San Diego Independent School District Follow Through project. San Antonio, Texas, Development Associates, 1977.

24. _____. Final report: ESAA evaluation, Edgewood High School. San Antonio, Texas, Development Associates, 1976.

25. _____. Final report: ESAA projects, 1975-1976 Edcouch-Elsa (Texas) Independent School District. San Antonio, Texas, Development Associates, 1976.

26. _____. Social attitudes of South Texas primary students. San Antonio, Texas, Development Associates, 1976.

27. _____ & WIEDING, JAYNE G. Report on spring 1976 testing, Comprehensive Tests of Basic Skills (Expanded Version), La Joya Indendependent School District, La Joya, Texas. San Antonio, Texas, Development Associates, 1976.

28. KENNEDY, JOHN J. The eta coefficient in complex ANOVA designs. Educational and Psychological Measurement (30):885-889, 1970.

29. LAMANA, RICHARD & SAMORA, JULIAN. Recent trends in educational status of Mexican-Americans in Texas. in ESTES, D. & DARLING, D. (eds.) Proceedings of the first Texas conference for the Mexican American. Austin, Texas, Southwest Educational Development Laboratory, 1967. pp. 20-41.

30. LOREA, LOUIS & KOHUT, SUZANNE. A comparative study of monolinguals and bilinguals in a verbal task performance. Journal of Clinical Psychology (17): 49-51. January 1961.

31. MAYESKE, GEORGE W. Educational achievement among Mexican-Americans. Washington, DC, National Center for Educational Statistics, 1969. (Technical Note 22)

32. MUNOZ, ADOLFO. Teacher-student values. in CERVANTES, ROBERT A. & JONES, EARL (eds.) Second interim report on the Edgewood School Plan. San Antonio, Texas, Development Associates, January 1976. pp. 96-102.

33. MYCUE, ELENA INEZ DE LOS SANTOS. Testing in Spanish and the subsequent measurement of English fluency. 1968. (ERIC ED 026 193)

34. ORTEGO, PHILLIP D. Montezuma's children. El Grito III(3):38-50. Spring 1970.

35. SHARP, MICHAEL; CERVANTES, ROBERT A. & JONES, EARL. A multidimensional pilot study of self-concept of Mexican American students. San Antonio, Texas, Development Associates, December 1975.

36. US COMMISSION ON CIVIL RIGHTS. Report II: the unfinished education. Washington, DC, Superintendent of Public Documents, October 1971.

37. _____. Report III: the excluded student. Washington, DC, Superintendent of Public Documents, May 1972.

38. _____. Report IV: Mexican American education in Texas - a function of wealth. Washington, DC, Superintendent of Public Documents, August 1972.

39. WELCH, RICHARD O. & JONES, EARL. Student tracking and attrition in ESP and comparison schools. in CERVANTES, ROBERT A. (ed.) Proceedings of technical seminar hosted by Development Associates and the National Institute of Education, Office of Experimental Schools. Rosslyn, Virginia, April 21, 1975.

APPENDICES A & B

SUMMARY TABLES OF SIGNIFICANT RESULTS:
1972-1973, 1973-1974

A1: Summary of Analysis of Variance: Arithmetic Subtest, Grade 8: 1972-1973*

Source	df	MS	F	p
Treatment	1	1144.06	4.33	.05
Error	58	264.51		
Total	59			

*Ns and means are shown in Table 2.

B1: Summary of Analyses of Variance, Kindergarten and Fifth Grade Classrooms: 1973-1974**

	Kindergarten: Language			
Source	df	MS	F	p
Treatment	5	13.34	5.49	.01
Error	12	2.43		
Total	17			
	Kindergarten: Social Studies			
Treatment	5	7.22	5.64	.01
Error	12	1.28		
Total	17			
	Fifth Grade: Language			
Treatment	5	53.08	2.95	.05
Error	17	17.99		
Total	22			

**Ns and means are included in Table 4.

B2: Summary of Analyses of Variance by Group (Treatment vs. Comparison) and by Grade (9th vs. 11th): 1973-1974*

CTBS/Q Test	Source	df	MS	F	p
Mathematics	Group	1	24651.90	113.04	.001
	Grade	1	26826.47	123.01	.001
	Interaction	1	1611.05	7.39	.01
	Residual	1293	218.09		
Language	Group	1	7182.01	47.29	.001
	Grade	1	36676.88	241.50	.001
	Interaction	1	118.46	.78	NS
	Residual	1373	151.87		
Reading	Group	1	6457.70	36.71	.001
	Grade	1	45047.50	256.08	.001
	Interaction	1	614.49	3.49	NS
	Residual	1379	175.91		
Total Battery	Group	1	96564.25	75.74	.001
	Grade	1	292602.94	229.50	.001
	Interaction	1	4479.88	3.51	NS
	Residual	1234	1274.99		
Study Skills	Group	1	4544.57	82.97	.001
	Grade	1	9284.71	169.52	.001
	Interaction	1	42.25	.77	NS
	Residual	1294	54.77		

*Ns and means are included in Table 4.

B3: ANCOVA Summary Tables for Changes in Treatment and Comparison Mean Scores: Grades 5→6, 10→11 (1972-1973→1973-1974)

	Grades 5 → 6				
CTBS/Q Test	Source	df	MS	F	p
Mathematics	Treatment	1	700.37	4.39	.05
	Error	293	159.62		
	Total	294			
	Grades 10 → 11				
Mathematics	Treatment	1	2174.39	6.78	.01
	Error	180	320.82		
	Total	181			
Total Battery	Treatment	1	8200.80	4.90	.05
	Error	176	1672.87		
	Total	177			

**Ns and means are shown in Table 5.

APPENDIX C

SUMMARY TABLES OF SIGNIFICANT RESULTS: 1974-1975

C1: Summary of Analysis of Variance by Group and Sex, Grade K: 1974-1975

TOBE Test	Source	df	MS	F	p
Language	Group	1	177.17	8.12	.01
	Sex	1	9.73	.45	NS
	Interaction	1	4.73	.22	NS
	Residual	292	21.81		
Mathematics	Group	1	508.59	28.61	.001
	Sex	1	2.02	.11	NS
	Interaction	1	8.99	.51	NS
	Residual	292	17.78		
Science	Group	1	314.27	13.99	.001
	Sex	1	136.26	6.07	.01
	Interaction	1	.00	.00	NS
	Residual	292	22.46		
Social Studies	Group	1	524.45	23.84	.001
	Sex	1	2.80	.13	NS
	Interaction	1	25.31	1.15	NS
	Residual	292	22.00		

		Treatment		Comparison		Total	
TOBE Test	Sex	$\bar{x}$	N	$\bar{x}$	N	$\bar{x}$	N
Language	M	21.31	65	20.05	65	20.68	130
	F	21.93	83	20.16	83	21.04	166
	Total	21.66	148	20.11	148	20.88	296
Mathematics	M	21.83	65	18.82	65	20.32	130
	F	21.31	83	19.00	83	20.16	166
	Total	21.54	148	18.92	148	20.23	296
Science	M	20.59	65	18.52	65	19.55	130
	F	19.22	83	17.16	83	18.19	166
	Total	19.82	148	17.76	148	18.79	296
Social Studies	M	20.68	65	17.35	65	19.02	130
	F	19.89	83	17.75	83	18.82	166
	Total	20.24	148	17.57	148	18.91	296

C2: Summary of Analysis of Variance by Group and Sex, Grade 1: 1974-1975

TOBE Test	Source	df	MS	F	p
Language	Group	1	173.07	10.79	.01
	Sex	1	87.72	5.47	.05
	Interaction	1	3.82	.24	NS
	Residual	344	16.04		
Mathematics	Group	1	114.05	6.61	.01
	Sex	1	17.06	.99	NS
	Interaction	1	6.22	.36	NS
	Residual	344	17.25		
Science	Group	1	372.74	14.48	.001
	Sex	1	17.39	.68	NS
	Interaction	1	8.26	.32	NS
	Residual	344	25.75		
Social Studies	Group	1	230.45	10.12	.01
	Sex	1	41.27	1.81	NS
	Interaction	1	11.46	.50	NS
	Residual	344	22.78		

		Treatment		Comparison		Total	
TOBE Test	Sex	$\bar{x}$	N	$\bar{x}$	N	$\bar{x}$	N
Language	M	20.87	97	19.25	84	20.12	181
	F	21.66	77	20.47	90	21.02	167
	Total	21.22	174	19.88	174	20.55	348
Mathematics	M	19.07	97	17.67	84	18.42	181
	F	19.25	77	19.25	90	18.78	167
	Total	19.15	174	18.03	174	18.59	348
Science	M	17.83	97	15.45	84	16.72	181
	F	17.07	77	15.31	90	16.12	167
	Total	17.49	174	15.38	174	16.43	348
Social Studies	M	19.35	97	17.37	84	18.43	181
	F	19.68	77	18.42	90	19.00	167
	Total	19.49	174	17.91	174	18.70	348

C3: Summary of Analysis of Variance by Group and Sex, Grade 2: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	242.01	1.28	NS
	Sex	1	223.95	1.18	NS
	Interaction	1	99.40	.53	NS
	Residual	312	189.07		
Language	Group	1	12.32	.09	NS
	Sex	1	89.10	.63	NS
	Interaction	1	1.70	.01	NS
	Residual	312	142.55		
Reading	Group	1	45.66	.40	NS
	Sex	1	96.75	.85	NS
	Interaction	1	22.49	.20	NS
	Residual	312	113.63		
Study Skills	Group	1	18.03	1.26	NS
	Sex	1	.33	.02	NS
	Interaction	1	6.08	.43	NS
	Residual	312	14.26		
Total Battery	Group	1	353.59	.34	NS
	Sex	1	18.58	.02	NS
	Interaction	1	256.59	.25	NS
	Residual	312	1043.17		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	38.53	80	35.71	87	37.06	167
	F	35.72	78	35.16	71	35.45	149
	Total	37.14	158	35.46	158	36.30	316
Language	M	26.81	80	27.07	87	26.95	167
	F	27.73	78	28.28	71	27.99	149
	Total	27.27	158	27.61	158	27.44	316
Reading	M	26.01	80	24.75	87	25.35	167
	F	26.59	78	26.39	71	26.50	149
	Total	26.30	158	25.49	158	25.89	316
Study Skills	M	9.83	80	9.61	87	9.71	167
	F	10.17	78	9.39	71	9.80	149
	Total	9.99	158	9.51	158	9.75	316
Total Battery	M	91.35	80	87.53	87	89.36	167
	F	90.04	78	89.83	71	89.94	149
	Total	90.70	158	88.56	158	89.63	316

C4: Summary of Analysis of Variance by Group and Sex, Grade 3: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	469.46	1.14	NS
	Sex	1	30.94	.08	NS
	Interaction	1	490.18	1.19	NS
	Residual	248	410.54		
Language	Group	1	347.82	1.25	NS
	Sex	1	75.45	.27	NS
	Interaction	1	324.23	1.16	NS
	Residual	248	279.42		
Reading	Group	1	417.33	1.72	NS
	Sex	1	.07	.00	NS
	Interaction	1	235.38	.97	NS
	Residual	248	243.31		
Study Skills	Group	1	162.91	5.99	.05
	Sex	1	.99	.04	NS
	Interaction	1	12.11	.45	NS
	Residual	248	27.20		
Total Battery	Group	1	303.17	.13	NS
	Sex	1	11.42	.01	NS
	Interaction	1	3079.00	1.29	NS
	Residual	248	2379.99		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	53.73	62	53.61	72	53.66	134
	F	55.80	64	50.07	54	53.18	118
	Total	54.78	126	52.10	126	53.44	252
Language	M	34.57	62	39.06	72	36.98	134
	F	37.92	64	37.85	54	37.89	118
	Total	36.27	126	38.54	126	37.41	252
Reading	M	31.42	62	35.82	72	33.78	134
	F	33.38	64	33.89	54	33.61	118
	Total	32.41	126	34.99	126	33.70	252
Study Skills	M	11.52	62	13.54	72	12.61	134
	F	12.08	64	13.22	54	12.60	118
	Total	11.80	126	13.41	126	12.60	252
Total Battery	M	119.71	62	128.49	72	124.43	134
	F	127.09	64	121.82	54	124.68	118
	Total	123.46	126	125.63	126	124.54	252

C5: Summary of Analysis of Variance by Group and Sex, Grade 4: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	1100.89	2.05	NS
	Sex	1	148.87	.28	NS
	Interaction	1	1313.71	2.44	NS
	Residual	326	537.57		
Language	Group	1	402.72	1.28	NS
	Sex	1	1675.80	5.32	.05
	Interaction	1	578.93	1.84	NS
	Residual	326	315.33		
Reading	Group	1	311.24	.90	NS
	Sex	1	.51	.00	NS
	Interaction	1	700.64	2.02	NS
	Residual	326	346.44		
Study Skills	Group	1	287.93	6.84	.01
	Sex	1	.85	.02	NS
	Interaction	1	58.24	1.38	NS
	Residual	326	42.09		
Total Battery	Group	1	5025.30	1.59	NS
	Sex	1	867.27	.28	NS
	Interaction	1	7530.07	2.39	NS
	Residual	326	3157.19		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	66.02	93	73.55	78	69.46	171
	F	68.71	72	68.22	87	68.44	159
	Total	67.19	165	70.74	165	68.97	330
Language	M	42.33	93	47.12	78	44.52	171
	F	49.54	72	49.00	87	49.25	159
	Total	45.48	165	48.11	165	46.79	330
Reading	M	40.81	93	45.58	78	42.98	171
	F	43.83	72	42.75	87	43.24	159
	Total	42.13	165	44.09	165	43.11	330
Study Skills	M	14.52	93	17.21	78	15.74	171
	F	15.26	72	16.26	87	15.81	159
	Total	14.84	165	16.71	165	15.78	330
Total Battery	M	149.16	93	166.24	78	156.95	171
	F	162.08	72	159.97	87	160.93	159
	Total	154.80	165	162.93	165	158.87	330

C6: Summary of Analysis of Variance by Group and Sex, Grade 5: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	685.94	1.99	NS
	Sex	1	46.94	.14	NS
	Interaction	1	187.61	.54	NS
	Residual	344	345.04		
Language	Group	1	2529.39	9.99	.01
	Sex	1	1278.20	5.05	.05
	Interaction	1	3.43	.01	NS
	Residual	344	253.33		
Reading	Group	1	564.39	2.47	NS
	Sex	1	25.50	.11	NS
	Interaction	1	10.53	.05	NS
	Residual	344	228.19		
Study Skills	Group	1	449.19	5.30	.05
	Sex	1	36.57	.43	NS
	Interaction	1	27.52	.32	NS
	Residual	344	84.81		
Total Battery	Group	1	10048.13	4.84	.05
	Sex	1	2270.82	1.09	NS
	Interaction	1	73.95	.04	NS
	Residual	344	2077.24		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	50.50	98	54.67	90	52.50	188
	F	52.72	76	53.94	84	53.36	160
	Total	51.47	174	54.32	174	52.89	348
Language	M	35.29	98	40.50	90	37.78	188
	F	38.93	76	44.55	84	41.88	160
	Total	36.88	174	42.45	174	39.67	348
Reading	M	36.57	98	38.80	90	37.64	188
	F	36.76	76	39.69	84	38.30	160
	Total	36.66	174	39.23	174	37.94	348
Study Skills	M	19.81	98	22.60	90	21.14	188
	F	21.03	76	22.69	84	21.90	160
	Total	20.34	174	22.64	174	21.49	348
Total Battery	M	122.36	98	133.97	90	127.92	188
	F	128.42	76	138.18	84	133.54	160
	Total	125.01	174	136.00	174	130.50	348

C7: Summary of Analysis of Variance by Group and Sex, Grade 6: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	190.58	.52	NS
	Sex	1	1826.29	4.93	.05
	Interaction	1	39.62	.11	NS
	Residual	374	370.35		
Language	Group	1	120.55	.41	NS
	Sex	1	4014.20	13.53	.001
	Interaction	1	459.83	1.55	NS
	Residual	374	296.70		
Reading	Group	1	236.85	.85	NS
	Sex	1	297.92	1.07	NS
	Interaction	1	13.67	.05	NS
	Residual	374	278.86		
Study Skills	Group	1	289.19	2.55	NS
	Sex	1	175.94	1.55	NS
	Interaction	1	30.80	.27	NS
	Residual	374	113.25		
Total Battery	Group	1	157.87	.06	NS
	Sex	1	15216.04	6.13	.05
	Interaction	1	988.25	.40	NS
	Residual	374	2483.92		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{x}$	N	$\bar{x}$	N	$\bar{x}$	N
Arithmetic	M	54.94	100	55.75	101	55.35	201
	F	58.70	89	60.81	88	59.75	177
	Total	56.71	189	58.11	189	57.41	378
Language	M	40.15	100	36.95	101	38.54	201
	F	44.47	89	45.69	88	45.08	177
	Total	42.19	189	41.02	189	41.60	378
Reading	M	40.95	100	39.01	101	39.98	201
	F	42.35	89	41.17	88	41.76	177
	Total	41.61	189	40.02	189	40.81	378
Study Skills	M	23.80	100	21.52	101	22.65	201
	F	24.60	89	23.46	88	24.03	177
	Total	24.18	189	22.42	189	23.30	378
Total Battery	M	136.04	100	131.71	101	133.87	201
	F	145.52	89	147.67	88	146.59	177
	Total	140.50	189	139.14	189	139.82	378

C8: Summary of Analysis of Variance by Group and Sex, Grade 7: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	106.06	.75	NS
	Sex	1	19.20	.14	NS
	Interaction	1	139.59	.98	NS
	Residual	478	142.11		
Language	Group	1	11.02	.07	NS
	Sex	1	4785.11	30.82	.001
	Interaction	1	1027.81	6.62	.01
	Residual	478	155.27		
Reading	Group	1	141.73	.98	NS
	Sex	1	.00	.00	NS
	Interaction	1	228.26	1.57	NS
	Residual	478	145.00		
Study Skills	Group	1	1.68	.03	NS
	Sex	1	732.47	13.10	.001
	Interaction	1	439.47	7.86	.01
	Residual	478	55.90		
Total Battery	Group	1	356.60	.34	NS
	Sex	1	5415.42	5.23	.05
	Interaction	1	3479.00	3.36	NS
	Residual	478	1035.89		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	36.63	112	34.60	128	35.55	240
	F	35.95	129	36.08	113	36.01	242
	Total	36.26	241	35.30	241	35.78	482
Language	M	29.65	112	27.02	128	28.25	240
	F	33.04	129	36.26	113	34.54	242
	Total	31.47	241	31.35	241	31.41	482
Reading	M	31.21	112	28.74	128	29.90	240
	F	29.84	129	30.12	113	29.97	242
	Total	30.48	241	29.39	241	29.93	482
Study Skills	M	17.43	112	15.63	128	16.47	240
	F	17.98	129	20.01	113	18.93	242
	Total	17.73	241	17.68	241	17.70	482
Total Battery	M	97.49	112	90.36	128	93.69	240
	F	98.82	129	102.46	113	100.52	242
	Total	98.20	241	96.03	241	97.12	482

C9: Summary of Analysis of Variance by Group and Sex, Grade 8: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	1429.67	7.33	.01
	Sex	1	834.80	4.28	.05
	Interaction	1	699.14	3.58	NS
	Residual	466	195.07		
Language	Group	1	335.19	1.82	NS
	Sex	1	4424.17	24.02	.001
	Interaction	1	1186.72	6.44	.05
	Residual	466	184.21		
Reading	Group	1	1819.11	8.77	.01
	Sex	1	22.77	.11	NS
	Interaction	1	767.02	3.70	NS
	Residual	466	207.43		
Study Skills	Group	1	569.11	8.48	.01
	Sex	1	484.59	7.22	.01
	Interaction	1	345.88	5.15	.05
	Residual	466	67.15		
Total Battery	Group	1	9755.55	6.84	.01
	Sex	1	10035.78	7.03	.01
	Interaction	1	7847.26	5.50	.05
	Residual	466	1427.31		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	42.52	98	36.10	94	39.38	192
	F	42.77	137	41.31	141	42.03	278
	Total	42.66	235	39.22	235	40.94	470
Language	M	35.66	98	30.15	94	32.96	192
	F	38.69	137	39.65	141	39.18	278
	Total	37.43	235	35.85	235	36.64	470
Reading	M	37.76	98	30.75	94	34.32	192
	F	35.62	137	33.81	141	34.70	278
	Total	36.51	235	32.58	235	34.55	470
Study Skills	M	21.25	98	16.98	94	19.16	192
	F	21.58	137	20.80	141	21.18	278
	Total	21.44	235	19.27	235	20.36	470
Total Battery	M	115.94	98	96.99	94	106.66	192
	F	117.08	137	114.76	141	115.90	278
	Total	116.60	235	107.65	235	112.13	470

C10: Summary of Analysis of Variance by Group and Sex, Grade 9: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	987.43	6.00	.05
	Sex	1	287.91	1.75	NS
	Interaction	1	233.59	1.42	NS
	Residual	780	164.56		
Language	Group	1	97.92	.68	NS
	Sex	1	6830.14	47.45	.001
	Interaction	1	17.89	.12	NS
	Residual	780	143.94		
Reading	Group	1	935.37	5.19	.05
	Sex	1	1688.95	9.37	.01
	Interaction	1	77.46	.43	NS
	Residual	780	180.17		
Study Skills	Group	1	167.98	2.87	NS
	Sex	1	195.93	3.34	NS
	Interaction	1	30.24	.52	NS
	Residual	780	58.61		
Total Battery	Group	1	2715.63	2.32	NS
	Sex	1	19799.08	16.88	.001
	Interaction	1	801.74	.68	NS
	Residual	780	1173.16		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{x}$	N	$\bar{x}$	N	$\bar{x}$	N
Arithmetic	M	38.21	194	37.04	203	37.62	397
	F	40.52	198	37.16	189	38.88	387
	Total	39.38	392	37.10	392	38.24	784
Language	M	31.99	194	33.00	203	32.50	397
	F	38.20	198	38.60	189	38.39	387
	Total	35.13	392	35.70	392	35.41	784
Reading	M	33.28	194	31.72	203	32.48	397
	F	36.85	198	34.03	189	35.47	387
	Total	35.08	392	32.83	392	33.96	784
Study Skills	M	19.80	194	19.27	203	19.53	397
	F	21.20	198	19.87	189	20.55	387
	Total	20.51	392	19.56	392	20.03	784
Total Battery	M	103.49	194	101.76	203	102.60	397
	F	115.56	198	109.79	189	112.74	387
	Total	109.58	392	105.63	392	107.61	784

C11: Summary of Analysis of Variance by Group and Sex, Grade 10: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	1018.08	4.38	.05
	Sex	1	41.58	.18	NS
	Interaction	1	5.73	.03	NS
	Residual	514	232.63		
Language	Group	1	2236.23	12.78	.001
	Sex	1	5724.91	32.72	.001
	Interaction	1	25.03	.14	NS
	Residual	514	175.00		
Reading	Group	1	1940.44	8.83	.01
	Sex	1	1925.55	8.76	.01
	Interaction	1	1.06	.01	NS
	Residual	514	219.73		
Study Skills	Group	1	899.45	13.17	.001
	Sex	1	206.82	3.03	NS
	Interaction	1	68.02	1.00	NS
	Residual	514	68.31		
Total Battery	Group	1	15189.69	10.11	.01
	Sex	1	12790.65	8.51	.01
	Interaction	1	70.98	.05	NS
	Residual	514	1502.28		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	42.67	114	40.08	134	41.27	248
	F	42.31	145	39.30	125	40.92	270
	Total	42.47	259	39.70	259	41.08	518
Language	M	37.99	114	34.28	134	35.99	248
	F	45.11	145	40.52	125	42.99	270
	Total	41.98	259	37.29	259	39.64	518
Reading	M	38.08	114	34.29	134	36.03	248
	F	42.04	145	38.07	125	40.20	270
	Total	40.30	259	36.12	259	38.21	518
Study Skills	M	22.40	114	18.99	134	20.56	248
	F	22.93	145	20.98	125	22.03	270
	Total	22.70	259	19.95	259	21.32	518
Total Battery	M	118.74	114	108.65	134	113.29	248
	F	129.46	145	117.89	125	124.10	270
	Total	124.74	259	113.11	259	118.93	518

C12: Summary of Analysis of Variance by Group and Sex, Grade 11: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	4206.16	16.36	.001
	Sex	1	33.70	.13	NS
	Interaction	1	570.46	2.22	NS
	Residual	352	257.18		
Language	Group	1	933.62	5.54	.05
	Sex	1	5588.65	33.17	.001
	Interaction	1	180.69	1.07	NS
	Residual	352	168.49		
Reading	Group	1	2424.90	11.29	.001
	Sex	1	1062.71	4.95	.05
	Interaction	1	113.81	.53	NS
	Residual	352	214.72		
Study Skills	Group	1	1737.77	25.27	.001
	Sex	1	151.15	2.20	NS
	Interaction	1	53.74	.78	NS
	Residual	352	68.78		
Total Battery	Group	1	20924.61	13.77	.001
	Sex	1	12805.49	8.43	.01
	Interaction	1	2303.53	1.52	NS
	Residual	352	1519.09		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	51.34	70	41.69	102	45.62	172
	F	49.36	108	44.86	76	47.50	184
	Total	50.14	178	43.04	178	46.59	356
Language	M	43.50	70	38.71	102	40.66	172
	F	50.09	108	48.20	76	49.31	184
	Total	47.50	178	42.76	178	45.13	356
Reading	M	43.90	70	37.40	102	39.84	172
	F	46.25	108	42.05	76	44.52	184
	Total	45.33	178	39.39	178	42.36	356
Study Skills	M	26.43	70	21.12	102	23.28	172
	F	26.95	108	23.22	76	25.41	184
	Total	26.75	178	22.02	178	24.38	356
Total Battery	M	138.74	70	117.79	102	126.32	172
	F	145.70	108	135.11	76	141.33	184
	Total	142.97	178	125.19	178	134.08	356

C13: Summary of Analysis of Variance by Group and Sex, Grade 12: 1974-1975

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	2674.12	7.79	.01
	Sex	1	10.43	.03	NS
	Interaction	1	82.33	.24	NS
	Residual	358	343.17		
Language	Group	1	726.41	4.02	.05
	Sex	1	1720.30	9.51	.01
	Interaction	1	60.54	.34	NS
	Residual	358	180.91		
Reading	Group	1	1188.89	4.40	.05
	Sex	1	82.58	.31	NS
	Interaction	1	20.91	.08	NS
	Residual	358	270.39		
Study Skills	Group	1	470.42	5.75	.05
	Sex	1	.79	.01	NS
	Interaction	1	.11	.00	NS
	Residual	358	81.86		
Total Battery	Group	1	12801.60	6.78	.01
	Sex	1	2240.51	1.19	NS
	Interaction	1	459.14	.24	NS
	Residual	358	1888.62		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	50.37	81	45.98	84	48.13	165
	F	50.99	100	44.68	97	47.88	197
	Total	50.71	181	45.28	181	48.00	362
Language	M	46.31	81	44.37	84	45.32	165
	F	51.51	100	47.93	97	49.75	197
	Total	49.18	181	46.28	181	47.73	362
Reading	M	47.46	81	44.36	84	45.88	165
	F	48.90	100	44.84	97	46.90	197
	Total	48.25	181	44.61	181	46.43	362
Study Skills	M	26.90	81	24.58	84	25.72	165
	F	26.96	100	24.71	97	25.85	197
	Total	26.93	181	24.65	181	25.79	362
Total Battery	M	144.14	81	134.70	84	139.33	165
	F	151.40	100	137.44	97	144.53	197
	Total	148.15	181	136.17	181	142.16	362

APPENDIX D

SUMMARY TABLES OF SIGNIFICANT RESULTS: 1975-1976

D1: Summary of Analysis of Variance by Group and Sex, Grade K: 1975-1976

TOBE Test	Source	df	MS	F	p
Language	Group	1	6.65	.35	NS
	Sex	1	76.77	4.05	.05
	Interaction	1	1.14	.06	NS
	Residual	228	18.98		
Mathematics	Group	1	196.26	10.15	.01
	Sex	1	6.27	.32	NS
	Interaction	1	4.79	.25	NS
	Residual	228	19.34		

		Treatment		Comparison		Total	
TOBE Test	Sex	$\bar{x}$	N	$\bar{x}$	N	$\bar{x}$	N
Language	M	20.52	56	20.72	65	20.63	121
	F	21.53	60	22.02	51	21.76	111
	Total	21.04	116	21.29	116	21.17	232
Mathemtics	M	20.00	56	18.43	65	19.16	121
	F	20.62	60	18.47	51	19.63	111
	Total	20.32	116	18.45	116	19.38	232

D2: Summary of Analysis of Variance by Group and Sex, Grade 1: 1975-1976

TOBE Test	Source	df	MS	F	p
Language	Group	1	2.96	.19	NS
	Sex	1	11.92	.76	NS
	Interaction	1	10.45	.66	NS
	Residual	352	15.78		
Mathematics	Group	1	261.53	14.57	.001
	Sex	1	.24	.01	NS
	Interaction	1	.57	.03	NS
	Residual	352	17.95		

		Treatment		Comparison		Total	
TOBE Test	Sex	$\bar{x}$	N	$\bar{x}$	N	$\bar{x}$	N
Language	M	21.32	90	20.81	94	21.06	184
	F	20.61	88	20.79	84	20.70	172
	Total	20.97	178	20.80	178	20.88	356
Mathematics	M	19.72	90	18.08	94	18.89	184
	F	19.75	88	17.95	84	18.87	172
	Total	19.74	178	18.02	178	18.88	356

D3: Summary of Analysis of Variance by Group and Sex, Grade 2: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	2455.93	13.78	.001
	Sex	1	466.06	2.62	NS
	Interaction	1	24.22	.14	NS
	Residual	246	178.25		
Language	Group	1	7025.38	42.91	.001
	Sex	1	1233.27	7.53	.01
	Interaction	1	57.58	.35	NS
	Residual	246	163.73		
Reading	Group	1	2728.30	21.32	.001
	Sex	1	165.17	1.29	NS
	Interaction	1	261.89	2.05	NS
	Residual	246	127.95		
Study Skills	Group	1	66.23	4.56	.05
	Sex	1	11.53	.79	NS
	Interaction	1	1.99	.14	NS
	Residual	246	14.54		
Total Battery	Group	1	34450.30	29.78	.001
	Sex	1	4838.32	4.18	.05
	Interaction	1	823.28	.71	NS
	Residual	246	1156.92		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	31.26	61	36.88	58	34.00	119
	F	33.38	64	40.24	67	36.89	131
	Total	32.34	125	38.68	125	35.51	250
Language	M	19.56	61	29.16	58	24.24	119
	F	23.05	64	34.57	67	28.94	131
	Total	21.34	125	32.06	125	26.70	250
Reading	M	21.90	61	23.36	58	24.08	119
	F	21.48	64	30.05	67	25.86	131
	Total	21.69	125	28.34	125	25.01	250
Study Skills	M	8.62	61	9.47	58	9.03	119
	F	8.88	64	10.08	67	9.49	131
	Total	8.75	125	9.79	125	9.27	250
Total Battery	M	72.72	61	92.40	58	82.31	119
	F	77.91	64	104.85	67	91.69	131
	Total	75.38	125	99.07	125	87.22	250

D4: Summary of Analysis of Variance by Group and Sex, Grade 3: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	1240.02	2.91	NS
	Sex	1	1222.56	2.86	NS
	Interaction	1	31.17	.07	NS
	Residual	234	426.83		
Language	Group	1	698.06	2.32	NS
	Sex	1	1255.11	4.18	.05
	Interaction	1	218.05	.73	NS
	Residual	234	300.36		
Reading	Group	1	159.39	.61	NS
	Sex	1	153.67	.59	NS
	Interaction	1	6.09	.02	NS
	Residual	234	261.27		
Study Skills	Group	1	27.38	.86	NS
	Sex	1	11.48	.36	NS
	Interaction	1	.04	.00	NS
	Residual	234	31.99		
Total Battery	Group	1	14.68	.01	NS
	Sex	1	6853.97	2.66	NS
	Interaction	1	319.68	.12	NS
	Residual	234	2573.08		

CTBS/Q Test		Treatment		Comparison		Total	
	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	54.52	66	50.62	61	52.65	127
	F	59.79	53	54.45	58	57.00	111
	Total	56.87	119	52.49	119	54.68	238
Language	M	35.27	66	40.49	61	37.78	127
	F	41.81	53	43.19	58	42.53	111
	Total	38.19	119	41.81	119	40.00	238
Reading	M	34.33	66	35.67	61	34.98	127
	F	35.62	53	37.60	58	36.66	111
	Total	34.91	119	36.61	119	35.76	238
Study Skills	M	12.38	66	13.03	61	12.69	127
	F	12.79	53	13.50	58	13.16	111
	Total	12.56	119	13.26	119	12.91	238
Total Battery	M	124.12	66	126.79	61	125.40	127
	F	137.23	53	135.24	58	136.19	111
	Total	129.96	119	130.91	119	130.43	238

D5: Summary of Analysis of Variance by Group and Sex, Grade 4: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	2372.64	4.68	.05
	Sex	1	51.75	.10	NS
	Interaction	1	182.29	.36	NS
	Residual	254	506.85		
Language	Group	1	2254.02	6.68	.01
	Sex	1	974.49	2.89	NS
	Interaction	1	125.65	.37	NS
	Residual	254	337.48		
Reading	Group	1	680.03	2.00	NS
	Sex	1	4.84	.01	NS
	Interaction	1	84.92	.25	NS
	Residual	254	339.94		
Study Skills	Group	1	126.98	3.41	NS
	Sex	1	17.09	.46	NS
	Interaction	1	58.56	1.57	NS
	Residual	254	37.28		
Total Battery	Group	1	14948.39	4.75	.05
	Sex	1	1311.25	.42	NS
	Interaction	1	1151.01	.37	NS
	Residual	254	3144.06		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	62.50	62	70.19	74	66.68	136
	F	65.08	67	69.38	55	67.02	122
	Total	63.84	129	69.85	129	66.84	258
Language	M	39.79	62	47.05	74	43.74	136
	F	45.09	67	49.55	55	47.10	122
	Total	42.54	129	48.12	129	45.33	258
Reading	M	39.58	62	43.93	74	41.95	136
	F	40.45	67	42.49	55	41.37	122
	Total	40.03	129	43.22	129	41.67	258
Study Skills	M	14.23	62	16.54	74	15.49	136
	F	14.66	67	15.06	55	14.84	122
	Total	14.45	129	15.91	129	15.18	258
Total Battery	M	141.87	62	161.18	74	152.38	136
	F	150.61	67	161.42	55	155.48	122
	Total	146.41	129	161.28	129	153.84	258

D6: Summary of Analysis of Variance by Group and Sex, Grade 5: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	567.00	1.85	NS
	Sex	1	1593.60	5.20	.05
	Interaction	1	1281.46	4.18	.05
	Residual	252	306.72		
Language	Group	1	2021.23	8.68	.01
	Sex	1	244.66	1.05	NS
	Interaction	1	59.32	.26	NS
	Residual	252	232.88		
Reading	Group	1	691.40	2.99	NS
	Sex	1	394.74	1.71	NS
	Interaction	1	225.57	.98	NS
	Residual	252	230.94		
Study Skills	Group	1	304.50	3.87	NS
	Sex	1	220.67	2.81	NS
	Interaction	1	220.13	2.80	NS
	Residual	252	78.60		
Total Battery	Group	1	9037.24	4.57	.05
	Sex	1	1948.90	.99	NS
	Interaction	1	3424.35	1.73	NS
	Residual	252	1978.42		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	50.70	77	57.82	62	53.88	139
	F	50.28	51	48.35	66	49.19	117
	Total	50.53	128	52.94	128	51.73	256
Language	M	37.33	77	43.87	62	40.25	139
	F	40.29	51	44.89	66	42.89	117
	Total	38.51	128	44.40	128	41.45	256
Reading	M	38.30	77	43.34	62	40.55	139
	F	37.73	51	38.97	66	38.43	117
	Total	38.07	128	41.09	128	39.58	256
Study Skills	M	20.61	77	24.52	62	22.35	139
	F	20.65	51	20.80	66	20.74	117
	Total	20.63	128	22.60	128	21.61	256
Total Battery	M	126.33	77	145.03	62	134.67	139
	F	128.29	51	132.21	66	130.50	117
	Total	127.11	128	138.42	128	132.76	256

D7: Summary of Analysis of Variance by Group and Sex, Grade 6: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	3214.53	10.30	.01
	Sex	1	3.42	.01	NS
	Interaction	1	885.13	2.84	NS
	Residual	286	312.01		
Language	Group	1	5963.47	22.17	.001
	Sex	1	727.71	2.71	NS
	Interaction	1	23.20	.09	NS
	Residual	286	269.04		
Reading	Group	1	3415.43	12.93	.001
	Sex	1	22.99	.09	NS
	Interaction	1	5.90	.02	NS
	Residual	286	264.08		
Study Skills	Group	1	1449.23	14.81	.001
	Sex	1	.31	.00	NS
	Interaction	1	312.44	3.19	NS
	Residual	286	97.85		
Total Battery	Group	1	37003.12	16.75	.001
	Sex	1	1130.28	.51	NS
	Interaction	1	1032.88	.47	NS
	Residual	286	2209.47		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	54.93	69	65.14	74	60.21	143
	F	58.65	76	61.86	71	60.20	147
	Total	56.88	145	63.53	145	60.20	290
Language	M	40.00	69	49.65	74	44.99	143
	F	43.74	76	52.25	71	47.85	147
	Total	41.96	145	50.92	145	46.44	290
Reading	M	40.84	69	47.42	74	44.25	143
	F	41.12	76	48.27	71	44.57	147
	Total	40.99	145	47.83	145	44.41	290
Study Skills	M	22.42	69	29.00	74	25.83	143
	F	24.43	76	26.86	71	25.61	147
	Total	23.48	145	28.95	145	25.71	290
Total Battery	M	135.77	69	162.20	74	149.45	143
	F	143.50	76	162.38	71	152.62	147
	Total	139.82	145	162.29	145	151.06	290

D8: Summary of Analysis of Variance by Group and Sex, Grade 7: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	152.06	.87	NS
	Sex	1	482.72	2.75	NS
	Interaction	1	8.20	.05	NS
	Residual	154	175.76		
Language	Group	1	31.01	.19	NS
	Sex	1	1027.32	6.28	.05
	Interaction	1	2.73	.02	NS
	Residual	154	163.58		
Reading	Group	1	1.83	.01	NS
	Sex	1	154.94	1.05	NS
	Interaction	1	23.58	.16	NS
	Residual	154	147.62		
Study Skills	Group	1	64.56	1.20	NS
	Sex	1	129.45	2.41	NS
	Interaction	1	31.53	.59	NS
	Residual	154	53.64		
Total Battery	Group	1	65.85	.06	NS
	Sex	1	4418.34	3.69	NS
	Interaction	1	36.81	.03	NS
	Residual	154	1198.54		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	X	N	X	N	X	N
Arithmetic	M	34.63	46	36.98	46	35.80	92
	F	38.64	33	40.06	33	39.35	66
	Total	36.30	79	38.27	79	37.28	158
Language	M	29.22	46	28.11	46	28.66	92
	F	34.12	33	33.55	33	33.83	66
	Total	31.27	79	30.38	79	30.82	158
Reading	M	26.24	46	27.11	46	26.67	92
	F	29.03	33	28.33	33	28.68	66
	Total	27.41	79	27.62	79	27.51	158
Study Skills	M	17.28	46	17.80	46	17.54	92
	F	18.21	33	20.55	33	19.38	66
	Total	17.67	79	18.95	79	18.31	158
Total Battery	M	90.09	46	92.20	46	91.14	92
	F	101.79	33	101.94	33	101.86	66
	Total	94.98	79	96.27	79	95.62	158

D9: Summary of Analysis of Variance by Group and Sex, Grade 8: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	1085.14	4.89	.05
	Sex	1	90.63	.41	NS
	Interaction	1	73.24	.33	NS
	Residual	334	222.09		
Language	Group	1	562.92	3.08	NS
	Sex	1	5333.65	29.20	.001
	Interaction	1	12.00	.07	NS
	Residual	334	182.67		
Reading	Group	1	4.42	.02	NS
	Sex	1	307.66	1.62	NS
	Interaction	1	32.10	.17	NS
	Residual	334	190.40		
Study Skills	Group	1	372.99	5.87	.05
	Sex	1	550.61	8.67	.01
	Interaction	1	115.02	1.81	NS
	Residual	334	63.50		
Total Battery	Group	1	2977.27	2.07	NS
	Sex	1	10018.41	6.96	.01
	Interaction	1	40.40	.03	NS
	Residual	334	1439.76		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	40.77	84	43.47	92	42.18	176
	F	40.88	85	45.44	77	43.05	162
	Total	40.83	169	44.37	169	42.60	338
Language	M	31.32	84	33.54	92	32.48	176
	F	38.91	85	41.88	77	40.32	162
	Total	35.14	169	37.34	169	36.24	338
Reading	M	33.20	84	33.57	92	33.39	176
	F	35.73	85	34.86	77	35.32	162
	Total	34.47	169	34.15	169	34.31	338
Study Skills	M	19.57	84	20.55	92	20.09	176
	F	20.97	85	24.29	77	22.54	162
	Total	20.27	169	22.25	169	21.26	338
Total Battery	M	105.30	84	110.58	92	108.06	176
	F	115.52	85	122.18	77	118.69	162
	Total	110.41	169	115.86	169	113.15	338

D10: Summary of Analysis of Variance by Group and Sex, Grade 9: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	131.55	.72	NS
	Sex	1	129.59	.71	NS
	Interaction	1	238.53	1.31	NS
	Residual	382	182.50		
Language	Group	1	389.84	2.57	NS
	Sex	1	2153.47	14.22	.001
	Interaction	1	70.49	.47	NS
	Residual	382	151.48		
Reading	Group	1	112.61	.76	NS
	Sex	1	207.33	1.39	NS
	Interaction	1	56.20	.38	NS
	Residual	382	148.70		
Study Skills	Group	1	14.47	.28	NS
	Sex	1	48.99	.95	NS
	Interaction	1	68.42	1.33	NS
	Residual	382	51.42		
Total Battery	Group	1	356.71	.31	NS
	Sex	1	5210.56	4.58	.05
	Interaction	1	982.01	.86	NS
	Residual	382	1138.97		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	39.17	94	36.18	73	37.86	167
	F	38.80	99	39.00	120	38.91	219
	Total	38.98	193	37.93	193	38.46	386
Language	M	32.50	94	33.53	73	32.95	167
	F	36.46	99	39.23	120	37.97	219
	Total	34.53	193	37.07	193	35.80	386
Reading	M	31.43	94	31.63	73	31.52	167
	F	32.16	99	33.92	120	33.12	219
	Total	31.80	193	33.05	193	32.43	386
Study Skills	M	20.50	94	19.14	73	19.90	167
	F	20.39	99	20.74	120	20.59	219
	Total	20.45	193	20.14	193	20.29	386
Total Battery	M	103.10	94	101.34	73	102.33	167
	F	107.41	99	112.14	120	110.01	219
	Total	105.31	193	108.06	193	106.68	386

D11: Summary of Analysis of Variance by Group and Sex, Grade 10: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	79.57	.33	NS
	Sex	1	8.20	.03	NS
	Interaction	1	199.29	.82	NS
	Residual	560	242.07		
Language	Group	1	28.19	.16	NS
	Sex	1	4236.69	23.50	.001
	Interaction	1	155.72	.86	NS
	Residual	560	180.32		
Reading	Group	1	940.08	4.15	.05
	Sex	1	841.84	3.72	NS
	Interaction	1	179.64	.79	NS
	Residual	560	226.30		
Study Skills	Group	1	592.07	8.47	.01
	Sex	1	33.54	.48	NS
	Interaction	1	4.26	.06	NS
	Residual	560	69.94		
Total Battery	Group	1	2014.92	1.28	NS
	Sex	1	8325.43	5.28	.05
	Interaction	1	1600.25	1.02	NS
	Residual	560	1575.59		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	44.31	137	44.75	145	44.54	282
	F	45.26	145	43.32	137	44.32	282
	Total	44.80	282	44.06	282	44.43	564
Language	M	38.05	137	38.66	145	38.36	282
	F	44.59	145	43.09	137	43.86	282
	Total	41.41	282	40.81	282	41.11	564
Reading	M	38.10	137	36.65	145	37.36	282
	F	41.68	145	37.96	137	39.87	282
	Total	39.94	282	37.29	282	38.61	564
Study Skills	M	25.06	137	22.83	145	23.92	282
	F	25.37	145	23.50	137	24.46	282
	Total	25.22	282	23.16	282	24.19	564
Total Battery	M	120.47	137	120.06	145	120.26	282
	F	131.52	145	124.37	137	128.05	282
	Total	126.15	282	122.15	282	124.15	564

D12: Summary of Analysis of Variance by Group and Sex, Grade 11: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	1275.18	4.68	.05
	Sex	1	141.99	.52	NS
	Interaction	1	441.90	1.62	NS
	Residual	300	272.27		
Language	Group	1	56.00	.34	NS
	Sex	1	2277.83	13.87	.001
	Interaction	1	75.85	.46	NS
	Residual	300	164.23		
Reading	Group	1	900.86	4.24	.05
	Sex	1	8.20	.04	NS
	Interaction	1	229.33	1.08	NS
	Residual	300	212.46		
Study Skills	Group	1	214.62	3.50	NS
	Sex	1	14.37	.23	NS
	Interaction	1	13.29	.22	NS
	Residual	300	61.31		
Total Battery	Group	1	5359.29	3.50	NS
	Sex	1	1495.64	.98	NS
	Interaction	1	2013.65	1.31	NS
	Residual	300	1533.43		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	X	N	X	N	X	N
Arithmetic	M	48.89	74	47.29	73	48.10	147
	F	49.94	78	43.51	79	46.70	157
	Total	49.43	152	45.32	152	47.38	304
Language	M	42.70	74	42.88	73	42.79	147
	F	49.18	78	47.35	79	48.26	157
	Total	46.03	152	45.20	152	45.62	304
Reading	M	43.72	74	42.07	73	42.90	147
	F	45.78	78	40.66	79	43.20	157
	Total	44.78	152	41.34	152	43.06	304
Study Skills	M	26.87	74	25.62	73	26.25	147
	F	27.72	78	25.63	79	26.67	157
	Total	27.30	152	25.63	152	26.46	304
Total Battery	M	135.31	74	132.23	73	133.78	147
	F	144.90	78	131.52	79	138.17	157
	Total	140.23	152	131.86	152	136.05	304

D13: Summary of Analysis of Variance by Group and Sex, Grade 12: 1975-1976

CTBS/Q Test	Source	df	MS	F	p
Arithmetic	Group	1	841.87	2.98	NS
	Sex	1	30.33	.11	NS
	Interaction	1	438.45	1.55	NS
	Residual	212	282.27		
Language	Group	1	109.67	.60	NS
	Sex	1	1965.40	10.74	.001
	Interaction	1	71.82	.39	NS
	Residual	212	182.99		
Reading	Group	1	1781.28	7.35	.01
	Sex	1	142.02	.59	NS
	Interaction	1	.29	.00	NS
	Residual	212	242.47		
Study Skills	Group	1	452.99	7.07	.01
	Sex	1	3.98	.06	NS
	Interaction	1	25.17	.39	NS
	Residual	212	64.07		
Total Battery	Group	1	6673.68	3.75	NS
	Sex	1	3813.91	2.14	NS
	Interaction	1	833.76	.47	NS
	Residual	212	1780.69		

		Treatment		Comparison		Total	
CTBS/Q Test	Sex	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N
Arithmetic	M	47.58	43	46.67	58	47.06	101
	F	51.28	65	44.60	50	48.37	115
	Total	49.81	108	45.71	108	47.76	216
Language	M	44.81	43	44.62	58	44.70	101
	F	52.11	65	49.58	50	51.01	115
	Total	49.20	108	46.92	108	48.06	216
Reading	M	46.47	43	40.59	58	43.09	101
	F	48.03	65	42.30	50	45.54	115
	Total	47.41	108	41.38	108	44.39	216
Study Skills	M	27.79	43	25.60	58	26.54	101
	F	28.77	65	25.20	50	27.22	115
	Total	28.38	108	25.42	108	26.90	216
Total Battery	M	138.86	43	131.88	58	134.85	101
	F	151.42	65	136.48	50	144.92	115
	Total	146.42	108	134.01	108	140.21	216